National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

NL-339 (r. 88/04) c

Canada

UNIVERSITY OF ALBERTA

SOCIAL DESIRABILITY, NONVERBAL CUES, AND INTERVIEWING METHOD

by

ALTA LYNN HICKEY

©

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF ARTS

SOCIOLOGY

EDMONTON, ALBERTA

FALL 1990

National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-64905-4

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR ALTA LYNN HICKEY

TITLE OF THESIS SOCIAL DESIRABILITY, NONVERBAL CUES, AND INTERVIEWING METHOD

DEGREE FOR WHICH THESIS WAS PRESENTED MASTER OF ARTS

YEAR THIS DEGREE GRANTED FALL 1990

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(SIGNED) Alta Lynn Hickey

PERMANENT ADDRESS:

129 Heathstone

Edmonton, Alberta

T6H 5E5

DATED August 20 1990

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled Social Desirability, Nonverbal Cues, and Interviewing Method submitted by Lynn Hickey in partial fulfilment of the requirements for the degree of MASTER OF ARTS.

.....
Harvey Krahn

.....
Michael Gillespie

.....
Adam Finn

Date..... June 7, 1980.....

Dedication

To more honest answers

Abstract

There is little consensus among researchers as to whether telephone or personal interviews produce better quality data. Many feel, though, that interviewing method can affect responses, and that this mode effect will be particularly great when the subject matter of questions is "sensitive." This thesis proposes that divergent opinions exist because little attention has been paid to understanding interview methods as communication situations involving nonverbal cues. Furthermore, by understanding "sensitivity" as a function of the social desirability of answering in a particular way, light may be shed on why nonverbal cues can cause differences.

Findings from the literature on mode effects, social desirability, sociolinguistics, and experiments on effects on task performance of communications medium are combined to develop and test a theory about circumstances under which nonverbal cues will affect answers to "sensitive" questions. It is proposed that the degree of social desirability involved in answering questions in a particular way is the critical determinant of whether nonverbal cues affect responses. Thus, evaluation cues should only be necessary information at certain levels of desirability. In face-to-face interviews, social desirability bias is hypothesized for responses to items with somewhat socially desirable or undesirable answers. Attention cues should produce less nonresponse overall, and better recall (where social desirability is not a factor).

Predictions were tested on responses to fourteen variables selected from the 1988 Edmonton Area Survey which used identical questions for a face-to-face (N=464) and a telephone (N=116) sample. Questions were chosen by asking experts and others to rate the social desirability of particular responses to particular questions, thus providing a range in subject matter social desirability.

Results indicate that social desirability-related mode effects are present, but they do not operate in quite the hypothesized ways. Norms about the social desirability of certain responses do create mode differences in answers, but the effect is limited to attitude questions with consensus that the answer is somewhat socially desirable or undesirable. Hypotheses were supported, but only on noncontroversial attitude questions.

Mode of interviewing appears to influence answers to behaviour/recall questions differently. Desirability of the response is less related to the question's content desirability than to the desirability of accommodative behaviour, shown by greater effort to provide detail in answering. This effect was surprisingly strong compared to that of item desirability.

Acknowledgments

So many people have played a role in helping me get this thesis written that it is impossible to thank them all here. A few, though, were especially helpful, and I would like to mention their contributions. My supervisor, Harvey Krahn, has been truly wonderful to work with. He always encouraged me to use my imagination freely, while almost imperceptibly nudging me to seek practical implications for my fanciful ideas. Harvey worked extremely hard with me on what turned out to be a very difficult topic, was prompt with his feedback, and always available for consultation - even from across the Atlantic while he was on sabbatical leave. I also appreciated his ability to sense exactly when to push me toward the next stage on the thesis.

My other committee members have also been very helpful. Mike Gillespie was always willing to offer me advice on how to bolster my statistical arsenal for testing my theory, while Adam Finn helped me narrow down my topic, and recommended useful sources. Other faculty members, Population Research Lab and Computing Services personnel have also contributed their time and advice by serving as my "panel of experts" for the social desirability assessment. For their input in this capacity, I would like to thank Harvey Krahn, Mike Gillespie, Adam Finn, Chuck Hobart, Les Kennedy, John Gartrell, Cliff Kinzel, and Chuck Humphrey. Many here-unnamed friends, neighbours, and Edmonton Area Survey interviewers also contributed as "nonexperts" in this same role.

I would also like to thank Linda Abbott for her invaluable aid deciphering Textform and "Thesis Layout" jargon, David Forde for showing me how to make SPSSX "do" logistic regressions, Dave Odynak for finding my survey data when they "disappeared", and Cliff Kinzel for filling me in on many details regarding how the survey was carried out. Jane Milliken, Frank Avakame, and Susan Shaw were all sympathetic listeners and informal advisors on problems with the thesis.

Finally, my family, Cliff, Clayton, and Travis deserve my gratitude for the encouragement and time they have given me, as well as my apologies for the neglect they have endured.

Table of Contents

Chapter — Page

I. Introduction 1

A. The Need for Explanatory Models 2

B. Organization of Thesis 5

II. Background Literature and Hypotheses 6

A. Telephones, Communication and Data Quality 7

B. Sensitivity of Questions 10

C. What Lies Behind Sensitivity 15

Social Desirability 18

Other Factors in Social Desirability 21

The Interviewer as a Factor 22

D. Role of Nonverbal Cues 26

E. Communications Medium Experiments 30

F. Hypotheses 32

III. Data and Methods 38

A. Data Set 38

B. Choice of Variables 40

C. Statistical Tests 45

IV. Results 47

A. Preliminary Findings - Item Desirability Hypotheses 47

Very Socially Desirable and Very Socially Undesirable 47

Somewhat Socially Desirable or Undesirable 50

Social Desirability Not a Factor 52

B. Variability 53

C. Nonresponse 56

V. Further Analysis 59

A. Item Desirability Hypotheses 60

Socially Desirable 60
Neutral Desirability 61
Socially Undesirable 64
Very Socially Undesirable 68
B. Variability 72
C. Nonresponse 73
VI. Conclusion 76
A. Summary of Mode Effects 76
Success of Predictions 77
Reevaluation of the Role of Nonverbal Cues 86
B. Place in the Literature 89
C. Directions for Further Research 93
References 97
Appendix One 105
Appendix Two 108

List of Tables

Table		Page
1	Predictions	34
2	Sample Differences on Means or Proportions	48
3	Variability in Responses of Face-to-Face and Telephone Samples	55
4	Proportion of Nonresponse for Face-to-Face and Telephone Respondents	58
5	Amount of ACCESS TV Watched Past Week by Mode and Level of Education	63
6	Means (by Mode) of High and Low Education Groups on Impaired Driver Punishments Question	70
7	Means (by Mode) of High and Low Education Groups on Driving After Drinking Question	71
8	Predictions	78
9	Summary of Findings	79

I. Introduction

This thesis asks whether we obtain different results in survey research if a questionnaire is administered to a respondent over the telephone rather than in person. But, more interestingly, it asks why there may be different results. Overall, the two methods have been found to yield very few major differences, but there have been a handful of consistent types of response differences noted. Such response differences are one type of "response effect" (Bradburn, 1983) deriving from the method or mode of administration of survey questionnaires.

Calling such response differences "mode effects" may be inappropriate since what is involved is really a comparison of two systems of data collection (Biemer, 1988) with design parameters and procedures which are only broadly equivalent. The term "mode effect" will still be used here but with the understanding that the primary effects cannot be thought of as part of a simple causal sequence, e.g., a telephone question produces "X" type of response. Thus, even such a basic issue as communicating a question is probably not the same process in the two interviewing conditions. Under such circumstances, a "mode effect" results from a whole complex of factors interracting differently with each other depending on the interviewing situation and the particular question asked.

Although response effects due to method of admininstration of survey questionnaires would also include those deriving from the use of Computer Assisted Telephone Interviewing techniques (CATI), as well as self-administration of questionnaires, these two techniques of data collection will not be discussed here. Because the thesis proposes and tests the idea that the cause of some response effects is due to differences in interpersonal communication between interviewer and respondents, consideration of effects is limited here to those in telephone and personal (face-to-face) interviews. These two methods of administering questionnaires differ only in the nature of the interaction occurring in the interview, whereas response effects due to use of self-administered questionnaires or CATI may arise from a host of other variables which cannot be easily controlled. Therefore, examining these other interviewing modes would preclude the possibility of testing ideas about the effects of

differences in communication on the types of responses people make.

The existence of response effects in face-to-face and telephone interviews has been the subject of much debate from at least as early as 1952 (Larsen, 1952). The primary concern has been whether the answer a person gives will be different if the question was asked by an interviewer who can be seen or by an interviewer who can only be heard over a telephone.

Underlying such discussions is a desire to know if telephone data are of the same quality as personal interview data since telephone interviewing has many obvious advantages. These include financial savings to the investigator, savings in time due to elimination of travel to respondents' residences and subsequent callbacks, as well as recently realized savings in data entry and analysis time with use of CATI. Other advantages include the increased control and the correction of interviewer effects this method can provide through monitoring of interviews by supervisors (Fowler, 1984). Furthermore, with many inner-city areas becoming increasingly inaccessible or even hostile to entry by personal interviewers, telephone surveys are seen as a useful alternative.

Thus, it has become increasingly important, for very practical reasons, to find out exactly how data obtained over the telephone differ in terms of both quality and quantity from data obtained from personal interviews.

A. The Need for Explanatory Models

Much of the work on differences between telephone and personal interviews has been limited to the discovery and enumeration of differences in response patterns between the two modes. This *ex post facto* type of approach has led to a wide range of often contradictory findings. For example, some studies report that telephone respondents are more willing to disclose embarrassing information, while others claim that respondents interviewed face-to-face will report such information more readily.

It is argued here that simply cataloging the differences which have been discovered will not lead to an overall understanding of the causal processes that produce differences in

answers. There must first be an understanding of how the two interviewing situations differ as social interaction and communication situations. With a more theoretical approach that focuses on differences in the kind of information that can be communicated, or on comprehension discrepancies between the two modes, hypotheses can be systematically tested, spurious explanations can be eliminated, and many of the contradictions might be resolved. The result should be a clearer understanding of why mode differences exist.

Groves (1987) implicitly recognizes the need for such an approach with his call for a theory of surveys which could incorporate explanatory findings from both cognitive psychology (relevant to the actual processing of questions) and from social psychology research relating to processes of compliance and influence. Such a theory would make use of what is known about the effects two speakers have on each other while communicating.

Others (e.g., de Leeuw and van der Zouwen, 1988) are beginning to develop very sophisticated graphic models for the explanation of mode effects. These models incorporate a wider range of interview mode characteristics and distinguish between more intervening variables than have previously been amalgamated. Such presentations suggest specific hypotheses to test on survey data because the logic behind expectations for particular findings is clearly laid out. For example, there can be several different dependent variables which describe performance characteristics of the mode of interviewing. These characteristics include response rate, amount of information, item response, accuracy of response, and absence of social desirability bias in response (Dillman, 1978). Having distinguished among them, we can then speculate about which of the various characteristics of the mode might affect one or more of these performance measures. Prediction of the effects of mode characteristics on data quality is facilitated by listing those conditions and intervening variables that would logically accompany a mode characteristic such as channel capacity (the kind and amount of information which can be conveyed). These other variables might include task complexity, motivation of respondents, need for approval, saliency of topic, and so on.

This thesis is concerned with some of those performance characteristics that derive from the mode characteristic of channel capacity. Because communication in personal

interviews is not limited to what is actually said, people in these situations experience a different kind of social interaction and information exchange than do those who are speaking over the telephone.

According to Reid (1977), visual cues should help explain differences in communication behaviour over the telephone. Hence, knowing the functions of such visual cues should enable one to predict differences in response patterns. For example, research on the functions of nonverbal signs has identified several types of signals which might influence a face-to-face respondent's answer. The signals demonstrating attention, comprehension, and evaluation have a "reaction function" by which the receiver [interviewer] provides feedback to the speaker [respondent] (Scherer, 1980).

This thesis develops and tests hypotheses about differences between telephone and personal interview response patterns which would be expected because of the presence or absence of these nonverbal communication signals. Specifically, the focus will be on what is usually referred to in the methodological literature as "sensitive" question response effects. The rationales for limiting the discussion are that: a) this area should be peculiarly subject to the effects of nonverbal feedback signals; b) findings on sensitivity are the most contradictory in the literature; and c) sensitivity usually involves peoples' notions of social desirability. A recent meta-analysis of data quality in mode comparison studies noted that absence of social desirability bias was one of the two indicators of data quality most affected by mode, the other being amount of information (de Leeuw and van der Zouwen, 1988). Thus, mode differences in social desirability bias appeared to be a promising area for investigation.

In general, the hypotheses presented at the end of the next chapter propose that the presence of evaluation, attention, and comprehension cues in face-to-face interviews will lead to more socially desirable responses, more uniformity in responses and fewer cases of nonresponse, as compared to telephone interviews.

Testing of these hypotheses will be carried out using data from the 1988 Edmonton Area Survey, a study designed to include both a telephone and a face-to-face sample. Thus, there were two randomly selected groups surveyed from the same population who differed

only by mode of questionnaire administration (Lalu, 1988). Questions used for the telephone survey were virtually the same as those used in the personal interview questionnaire (with a few modifications to some response categories). Hence, any differences in response patterns between the groups is probably a result of differences in interviewing method.

The fact that this survey is a local one, carried out every year by the Population Research Laboratory of the Sociology Department of the University of Alberta, also made this data set a good choice because it was possible to ask questions about how the survey was carried out. Being able to talk to some of the interviewers who worked on the survey was also of benefit. Finally, the fact there were a fair number of questions asked which had socially desirable or undesirable answers made the Edmonton Area Survey a useful body of data for this thesis.

B. Organization of Thesis

Chapter two reviews the literature on "sensitive questions" and how they relate to various conceptions of social desirability. It also considers studies of nonverbal communication and experiments in the social psychology of telecommunications which examine the effect on task performance of seeing, rather than merely hearing, the person with whom one is communicating. Ideas from these areas are then used to provide support for several hypotheses about expected mode-related differences in survey responses.

The third chapter describes the data set and outlines the methods used to select questions for analysis. There is also a discussion of the statistical procedures employed and the criteria for making decisions about the existence of a mode effect.

Chapter four presents the initial results of the analysis, while Chapter five describes some further analysis clarifying previous findings and unexpected results.

The last chapter draws conclusions about the extent to which the hypotheses were supported and offers some thoughts concerning the value of carrying out this type of study.

II. Background Literature and Hypotheses

Most writings on the subject of response effects see their origin in three primary sources, the respondent, the interviewer, and the characteristics of the task itself, with the variables affecting the task being the most important (Bradburn, 1983). Task variables themselves have been divided into five types by Bradburn: a) method of administration; b) open versus closed questions; c) question order; d) length and wording of questions; and e) memory.

For this study, the concern is with method of administration, but the effect of this variable cannot be thought of as independent of the other task variables, or of either respondent or interviewer effects, for that matter. Such interdependence becomes obvious, for example, in the case of an interaction between method and respondent characteristics whereby the effect of interviewing by telephone might be greater for an older respondent than it is for a younger respondent.

The usual concern in studies of the effects of mode of administration is whether the data obtained in different modes are of the same quality, and, consequently, comparable. In other words, what is the amount of "error" in response introduced by use of one or another mode? If "quality" is being used as the sole indicator of difference, the personal interview has been declared "superior" in an one recent assessment which found such interviews to be "better" on thirteen of twenty-five standards (Smith, 1984). But bases for comparison between studies have not always been uniform, and, even when they have been, conclusions have differed. Furthermore, most studies of data quality have simply assumed that data gathered in the face-to-face mode set the standard by which other methods must be judged.

Bradburn's (1983) review of previous studies (Sudman and Bradburn, 1974; Dillman, 1978; and Groves and Kahn, 1979) found "no consistent or large effects" for method. There appears to be a slight amount of overreporting for behavioural items in personal interviews compared to a slight underreporting of these items in telephone interviews. However, in general, results are not consistent and would seem to be a function of other factors such as the sensitivity of the questions asked (Bradburn, 1983).

A more recent meta-analysis (deLeeuw and van der Zouwen, 1988) which reviewed twenty-five such studies comparing face-to-face and telephone responses has concluded that two areas of data quality appear to be most strongly affected by mode. These are social desirability bias (with "quality" measured by its absence), and the amount of information obtained. Face-to-face interviews fared better on both indicators of data quality. That is, they obtained more information and, surprisingly, exhibited less social desirability bias.[1] The three other data quality indicators examined were accuracy, item (non)response, and similarity (of response distributions).

Thus, this work is probably the best assessment so far of those particular types of mode effects because it attempts to evaluate the relative strength of these effects on different indicators of data quality. But there are still difficulties drawing conclusions because of the small number of studies which measure the same dependent variable assessing data quality. The authors derive some reassurance, though, from the fact that even for the most affected areas of data quality, mode effect sizes were relatively small.

A. Telephones, Communication and Data Quality

In general, peoples' experience with telephones and their effects on our lives have received scant attention in the literature on data quality. But ignoring the role of the telephone may be hazardous to our understanding of mode effects. Compared to the work on radio, television, and newspapers, the effect of telephones has received little attention (Keller, 1977). However, there have been some attempts to document ways the telephone has changed us and how we have changed our usage of the telephone.[2] It is entirely conceivable that such knowledge may ultimately be necessary for any real understanding of why mode effects in interviewing can occur. But, for the present, only a few of these ideas, those most closely related to the thesis topic, can be mentioned.

---

[1] Only three studies, though, could be compared on this social desirability bias indicator.

[2] For example see McLuhan (1964) Ball (1968) Aronson (1971, 1977) Cherry (1977) Keller (1977) Boettinger (1977) and Goodman (1982).

The perceived function of the telephone appears to have changed drastically, from being merely a way to send orders to becoming a tool having "conversational power" (Cherry, 1977:119). This is a change having profound implications for the nature of our daily existence. It has been responsible for changes in the conduct of business, police work, crime, newsgathering and reporting, institutional shifts in education, medicine, law and warfare, as well as changes in our personal lives with respect to manners, morals, how we deal with crises, leisure activities, ordinary life routines, and even in how families adapt to modernization (Aronson, 1971). We may feel that telephones are simply a communication device that we use to accomplish our goals. But telephones appear to have an independent effect on our lives which must be due, in part, to the fact that communication using them differs in kind from communication we engage in face-to-face.

Such ideas have not really been explored in the literature on factors in data quality, and they cannot be more than mentioned here. Nevertheless, the evidence of a decrease over time in the effect size of interviewing mode (de Leeuw and van der Zouwen, 1988) may well suggest something about a change in the nature of telephone communication itself. Telephone interactions may have become more like face-to-face interactions for many people.

One major change brought about by telephones is a democratizing of social relationships because of the impossibility of communicating signs and symbols of status over the telephone (Ball, 1968). This type of trend could very well have implications for the amount of social desirability bias in answers which people would give to telephone interviewers. If relative status cannot be perceived, respondents may be less likely to attempt to impress interviewers with desirable responses. But if the telephone has worked to democratize social relationships generally, we might then expect a decrease in social desirability bias over time, regardless of whether the interview is conducted in person or over the telephone.

Another relevant topic may be the sense of freedom to discuss particular topics or to say things in ways not usually tolerated in face-to-face interaction for which the telephone may be responsible. Aronson cites early commentators (from the 1890's and 1900's) on the

harmful effects of the phone. They objected to the "new codes of conduct" and the "absence of inhibition" it produced, as in the case of the impulsive women who "say things to men and to each other over the telephone that they would never say face to face" (1971:165). Without comparative studies, though, it is impossible to estimate how much such changes can be attributed to the telephone *per se*, or simply to the telephone's unique effect in what was a strait-laced Victorian society (Aronson:166).

However, this perspective could still prove useful in understanding particular mode effects noted by some researchers. For example, the tendency for more "extremeness" response bias (choice of extreme answer category) in telephone respondents has been noted by Jordan *et al.* (1980). This may be partially explainable with reference to the use of telephones to convey extreme versions of ideas and opinions unacceptable in face-to-face interaction. Such explanations might be considered in addition to more conventional ones for "extremeness", such as the likelihood that scale endpoints would become more salient to the respondent if answer categories must be read aloud by an interviewer.

There are certainly other mode effects which have been noted by researchers which may be reinterpreted. For example, the "optimism" of responses over the telephone (Groves and Kahn, 1979) may have more to do with "proper" telephone etiquette than with the actual topic of conversation and how it interacts with mode. Or, the often-noted ability of telephone interviews to elicit more information on topics considered embarrassing or threatening may have something to do with the acceptability of disclosing such information by telephone caused by the proliferation of telephone self-help services and crisis lines.

A complicating factor in the matter of sensitive questions and telephones may be the peculiar composition of some samples. The conclusions drawn by Sykes and Collins (1987) about whether information about sensitive subjects can be obtained by telephone, for example, are subject to whatever limitations are imposed by a telephone ownership rate of only 75% in the United Kingdom. Not only could the "sensitivity" of some items by telephone be different if telephones are owned primarily by the wealthier people in a society, but there may be a very different kind of communication pattern for telephone conversations where

phones are comparatively rare."

These influences on sensitivity could derive from a society's unique experience with telephones. In addition, any society would have its own set of issues which were uniquely sensitive. Such topics have not been explored in the data quality literature.

The different ways in which the telephone affects us are relevant to understanding why there may or may not be mode effects on survey responses for one or another population or topic. But examining this area in detail would constitute another thesis. Therefore, the following discussion is restricted to findings regarding relevant factors which likely underlie the apparent mode effect on social desirability bias.

It should be noted, though, that de Leeuw and van der Zouwen (1988) found significant differences between telephone and personal interviews on all indicators of data quality except accuracy, with the largest effect sizes of mode on absence of social desirability bias and amount of information. Their meta-analysis revealed less social desirability bias for personal interviews. This was a finding which they found surprising, but it must be remembered that it was based on a sample of only three studies.[3] Further examination of this topic is clearly necessary.

B. Sensitivity of Questions

Mode of interviewing has been found to affect both the way a respondent answers questions on sensitive topics as well as his/her willingness to answer such questions at all (Sudman and Bradburn, 1974). Those who have reviewed studies on the subject, though, have reached opposite conclusions as to how mode affects answers to sensitive questions, as witnessed by the debate about whether more or less information is given over the telephone. For example, Sykes and Collins (1987) feel that people are more willing to answer sensitive questions on the telephone than they are in a face-to-face situation. Reasons given for this

[3]de Leeuw and van der Zouwen (1988) do not report sample sizes for each of the studies included in their meta-analysis. They state, though, that the number of respondents varied widely and that the average sample size was 1532. Actual N's for the three studies using social desirability bias as a data quality indicator were 1600 (Colombotos, 1969), 1167 (Henson *et al.*, 1978), and 1239 (Herman, 1977).

finding usually emphasize the relative sense of anonymity provided by telephone interviews in comparison with personal interviews. Major differences in responses to sensitive questions are, thus, seen to be a function of the degree of impersonality of the interview situation. It follows that a self-administered questionnaire would, theoretically, elicit the most open, honest answers. Face-to-face interviews should generate the least, and responses to telephone-administered questions would be assumed to lie between the two extremes (Bradburn, 1983).

Other reports, however, recount that there is less willingness to answer sensitive questions over the telephone. Fowler's (1984) summation of studies indicated a stronger social desirability bias in telephone responses. Frey's (1983) assessment was that there had been a shift over time from less to more item nonresponse (as an indicator of sensitivity) in telephone surveys.

One source of such differences may simply be the fact that different issues, which need not be equally sensitive, were examined in different studies. For example, Hochstim's (1967) finding of more admissions of alcohol consumption over the telephone has been used to argue that people are more willing to discuss sensitive subjects on the telephone. Others feel that telephones deter discussion of sensitive subjects, given findings such as those of Cannell, Oksenberg, and Converse (1979) that there are lower levels of health symptom reporting on the telephone.

It is not necessary to review all of the studies which have compared telephone and face-to-face responses to sensitive questions. Reviews of this sort already exist[4] and, as noted, they have reached different conclusions. Part of the problem lies in the fact that not every review has examined the same studies. But there are also cases where the same findings have been interpreted differently by the reviewers. For example, de Leeuw and van der Zouwen (1988) state that Hochstim (1967) reported that data collection strategies were practically interchangeable in asking questions about general health. But Bradburn (1983) and Biemer (1988) assert that Hochstim, in the same study, found a greater tendency for telephone

[4]See, for example, de Leeuw and van der Zouwen (1988) Fowler (1984) Frey (1983) Sudman (1983) Groves and Kahn (1979) Sudman and Bradburn (1974).

respondents to report negative information about themselves. Thus, how a study's findings are interpreted may depend on such factors as the purpose of the review and on the level of detail in findings which is being examined.

Even when exactly the same sensitive subject is examined by reviewers, there may be differing assessments of the same mode effects studies. De Leeuw and van der Zouwen (1988), for example, cite Mangione *et al.*'s (1982) study on alcohol consumption and report that equivalent results were obtained in both modes. Fowler (1984), however, claims that this study indicates that telephones produce a social desirability bias because there were fewer reports of past drinking problems on the telephone. My own reading of the study is that this single difference in reporting of drinking problems was the only one found, and that one should be hesitant, then, to conclude that a "mode effect" exists. Further, it may be risky to attribute the difference in reporting to a social desirability bias when other factors have not been ruled out, such as a difference in recall effort between telephone and face-to-face respondents.

Many reviews of mode comparison studies have stressed the breadth of the range of sensitive topics for which data can be obtained as validly or as readily by telephone as by personal interview.[5] These topics include alcohol consumption (Mangione *et al.*, 1982; Sykes and Collins, 1988), health/depression (Aneshensel *et al.*, 1982), health/opinion/attitudes (Colombotos, 1969), health/cervical cytology (Hochstim, 1967), health/self disclosure (Janofsky, 1971), health/general (Hochstim, 1967), crime/victimization (Klecka and Tuchfarber, 1978), and personal facts (Bradburn and Sudman, 1979).

Other studies, by contrast, have found the telephone to provide <u>less</u> valid or less readily obtainable data concerning the supposedly sensitive topics of health/moods (Henson *et al.*, 1978), voting behaviour (Herman, 1977), and, especially, income (Rogers, 1976; Groves and Kahn, 1979; Jordan *et al.*, 1980). However, an exhaustive recent study of income reporting in Denmark (Körmendi, 1988) revealed more nonresponse for telephone

[5]Most assessments and subject matter designations of studies named here are taken from a list provided by de Leeuw and van der Zouwen (1988), but I have used my own judgment in characterizing subjects as "sensitive".

respondents, albeit equally accurate (or inaccurate) reporting, in both modes (i.e., slight overreporting face-to-face, slight underreporting on the telephone).

Sometimes conclusions have been drawn that telephone interviews can successfully obtain data about sensitive subjects, such as income or level of education, when the telephone interview was actually a reinterview after a previous personal interview with the same respondents (Rogers, 1976). In another study, Coombs and Freedman felt that data could be successfully obtained by telephone about the sensitive topics of family planning and fertility. Again, the telephone interview had followed a personal interview (Coombs and Freedman, 1964). In such cases, it may not be possible to draw valid conclusions about effects of the telephone because of the possible contaminating effects of previous interviewing experience. There is evidence, too, from other studies that in the case of reinterviews, attitudes are less consistent when the initial interview is in person and the follow-up by telephone than they are when both interviews are of the same type (Herman, 1977). Thus, there may be problems in drawing conclusions about mode effects on reporting of sensitive subjects when evidence has been based on reinterviews.

There are also studies which simply report on the ability of telephone interviews to elicit information on sensitive topics without actually comparing their performance with personal interviews using the same sampling population and questionnaire. For example, Smith (1985, 1986) was able to elicit information about woman abuse using telephone interviews which revealed abuse rates comparable to those obtained in other studies using face-to-face interviews. He felt that telephone interview methods were probably superior for such sensitive topics because of the anonymity, confidentiality, and control over interviewer effects which they provide.

Despite the existence of some contrary opinion, including the recent meta-analysis (de Leeuw and van der Zouwen, 1988), it seems that the majority of literature reviews have concluded overall that telephone interviews can be superior in obtaining information about many subjects that should be sensitive, and probably can do this because of the sense of anonymity they offer. There is at least one study, however, which calls into question the

supposed facilitating effect of anonymity on disclosure. Wildman (1977) concluded that anonymity's benefits may be limited to specific topics when he tested the effects of anonymity and setting on responses to a self-administered questionnaire. The subject matter, formation of teacher unions, was assumed to be a threatening one to teachers if the questionnaire was received at work and/or if an identification number on the return postcard seemed to belie the promised anonymity. Neither setting nor anonynmity nor both together affected response, however.

While most researchers feel that the telephone's anonymity increases response to sensitive items, the relationship may not be so straightforward. As Bradburn (1983:298) puts it, the more anonymous methods of data collection appear to "lower the degree of under or overreporting", but that "the data here are not entirely consistent and more work needs to be done to define precisely those conditions under which the more anonymous methods work better."

Consistency may be elusive because of various methodological problems with many of the studies which can cast doubt on findings of mode effects. For example, as has been mentioned, when findings are based on results from reinterviews, especially reinterviews using a different interviewing method, causes of effects cannot be so readily attributed to mode. In other cases, studies have compared results of interviews conducted by entirely different types of sponsoring organizations. The importance which Dillman (1978) places on trust for obtaining valid answers may also extend to trust of the organization sponsoring the survey as well, and this type of factor may even outweigh the benefits of the anonymity provided by telephones.

Both of the previous examples illustrate problems caused by drawing conclusions when other alternative causal explanations have not been eliminated. Of course, eliminating alternatives without running actual experiments would be difficult. Nevertheless, competing explanations such as differential recall in the two interviewing modes, must at least be considered. Such a course is preferable to simply making the assumption that sensitivity to mode has been demonstrated because there was a difference by mode in some measure of

performance characteristics. Certainly, some of the studies which look at reporting of health symptoms in both modes, and conclude that the telephone produces a social desirability bias because fewer moods and symptoms are reported (e.g., de Leeuw and van der Zouwen's, 1988 interpretation of findings in Henson *et al.*, 1978), could be reinterpreted to show that people may have better recall in personal interviews, or that they put more effort into remembering when asked to do so face-to-face.

It may also be necessary to distinguish between mode's effect on behavioural as opposed to attitudinal items. Sudman and Bradburn (1974) found that while telephone administration of surveys had more of an overall response effect than face-to-face administration, the effect was more pronounced for attitudinal items than for behavioural items. Since their examination of this topic was not restricted to "sensitive" attitudes and behaviours, we cannot be certain whether telephone administration would have the same augmented effect on answers to sensitive attitudinal questions that it has for other attitudes.

C. What Lies Behind Sensitivity

The greatest problem in these mode comparison studies on sensitive subjects, though, is that conclusions have been drawn without first really establishing the sensitivity of a subject. Usually, there has been an asssumption on the part of the researcher that a topic is sensitive, that is, that it creates a situation in which the respondent is motivated to present him or herself favourably or to reduce anxiety by the way he or she responds. Response effects based on better self-presentation would be found with: 1) questions that threaten the respondent and arouse anxiety; 2) questions asking for information on topics that have very desirable answers; 3) questions to which the respondent wants to appear to know the answer when he is, in fact, ignorant[6]; and 4) attitude questions for which norms of politeness operate to produce acquiescence to the interviewer's opinion or to the social situation of the interview.

[6]Some recent research, though, has questioned the role of social desirability, suggesting that responses to fictitious issue questions are more a function of pressure created by the way a question is asked or by how "don't know" responses are handled by the interviewer (Bishop *et al.*, 1986).

with the aim of avoiding conflict (Sudman and Bradburn, 1974).

Researchers frequently assume that a topic is "sensitive" in one of the above ways without establishing that this is the case, and then conclude that mode of interviewing does or does not affect peoples' answers on such a sensitive subject. For example, the Coombs and Freedman (1964) investigation involved asking women questions over the telephone about the supposedly sensitive subject of fertility. However, their response rate of 97.6% and the reported eagerness of many respondents to participate would seem to belie the sensitivity of the subject, at least for their sample.

Subjects might also be "sensitive" in peculiar ways, as in the case of reporting income. Here, sensitivity may have less to do with self-presentation than with peoples' experiences of being deluged with offers over the phone from various businesses after having made the mistake of reporting their incomes. Thus, sensitivity of income information may have a very different basis than the self-presentation issue behind the sensitivity of other topics. In order to claim that mode of interviewing affects how people answer questions on sensitive subjects, it would first be desirable to establish the fact that the subject is "sensitive" and in what way it is so.

Related to this idea of establishing that a subject is sensitive is the possibility that the degree of sensitivity can vary by topic. Many studies make the implicit assumption that topics are equally sensitive and that mode, then, does or does not have an effect on responses to "sensitive" questions. An important exception is a study carried out by Locander, Sudman, and Bradburn (described in Bradburn and Sudman, 1979) in which the degree of threat of the question was varied, and the effect of threat on response distortion was checked for four different interviewing methods (face-to-face, telephone, self-administration, and random response methods). Questions tested concerned behaviours which could be checked for accuracy and included the following behaviours, in order of presumed threat level: having a public library card; voter registration; voting; involvement with bankruptcy; and being charged with drunken driving.

As predicted, response distortion increased with level of threat, but it appears that in drawing this conclusion, the researchers used the level of distortion as a measure of level of threat itself, which may not be a valid procedure. Thus, on the basis of the amount of distortion discovered, they reordered the items on threat level so that bankruptcy was now considered to be less threatening than voting behaviour. A better procedure would have entailed a prior, independent ranking of threat level for the items. They also found weak evidence of overreporting of socially desirable acts with the more personal methods and underreporting of socially undesirable acts for the more anonymous methods. However, there really were no substantial differences in distortion between telephone and personal interviews.[7]

Another way that degree of sensitivity has been examined, though in an indirect way, is in a study by Henson *et al.* (1978) on mode differences in reporting moods, health symptoms, and need for social approval. Three scales were used, the Lubin Depression Scale, the Midtown Symptom Scale, and the Marlowe-Crowne Scale. Telephone respondents showed higher means (more socially desirable responses) on all three. When the authors looked at the results more closely, however, they observed that respondents were more willing to make disclosures in the personal interviews for only some items in the scales. For example, disclosures were actually being made on the less private items of the depression and symptom subscales. They hypothesized, then, that mode was having an effect for items of intermediate privacy, where the personal interview led to greater reporting, but they could offer no explanation for this finding.

Thus, there is a fair amount of evidence indicating that the issue of how mode is related to the sensitivity of a topic has not been satisfactorily dealt with in the literature. The existence of very contradictory findings, including the meta-analysis result indicating less social desirability bias in personal interviews, are symptoms of two problems; the

---

[7]Conclusions drawn might be questioned because of the fact that the degree of threat may not have been judged to be the same by respondents and researchers (e.g., in the bankruptcy question) and that different, non-random samples were used for each question (e.g., a sample of all those convicted of drunk driving in the past six months).

conceptualization of sensitivity, and how it relates to the interpersonal dynamics of interviewing situations.

Furthermore, there is reason to believe that the degree of threat of a topic does make a difference in some ways. For example, the study described in Bradburn and Sudman (1979) found that the degree of threat influences the amount of distortion, but that there was no difference by mode (telephone and face-to-face). However, in the Henson *et al.* (1978) study, a mode difference was found for a different topic, moods, health symptoms, and need for social approval. Here, though, mode's effect was mediated by the degree of privacy of the item's content.

These studies indicate that items in surveys differ in degree of sensitivity, that responses are affected by this differentiation, and that at least for some topics, mode differences may be related to an item's degree of sensitivity. This sort of differentiation in degree of sensitivity of an item forms the basis for the hypotheses which are presented at the end of this chapter. Basically, the degree of social desirability of different items should vary, and because of this variation, nonverbal communication ought to play a greater role at certain levels of this desirability. For the purposes of this thesis, then, sensitivity of items is assumed to be rooted in the social desirability of the items. The distinction made by Sudman and Bradburn (1974) between sensitivity based on threat to the respondent and that based on the social desirability of the topic is considered unnecessary. The other two types of sensitivity mentioned by these authors, those based on a desire to appear knowledgeable or to acquiesce to the interviewer's opinion, will not be specifically addressed here.

Social Desirability

The whole issue of the source of social desirability as a motivating force in responses to survey questions is one fraught with debate, as De Maio (1984) has pointed out in her review of the literature. Even her seemingly simple definition of the concept of social desirability, "a tendency to give a favorable picture of oneself" (p. 257), taken from Sellitz, Wrightsman, and Cook (1976) immediately raises questions as to the circumstances under

which this tendency operates and about who decides what is "favorable". All definitions of social desirability, however, do share the underlying assumptions that some things are "good" while others are "bad", and that respondents generally want to answer questions in a way that will make them appear good.

There are three areas in the literature that are plagued by inconsistencies and debate: "(1)Does social desirability refer to a personality characteristic or an item characteristic? (2)What does 'desirability' include? (3)Who sets the standard for determining desirability?" (De Maio, 1984:258). The personality/item characteristic debate is an argument between those who see social desirability as a personality characteristic (e.g., an individual need to conform to social standards, or need for social approval), and those who feel that there are various behaviours and opinions which are more or less socially desirable. The latter would argue that there are norms concerning these opinions and behaviours. Hence, social desirability must be discussed in relation to particular survey items.

Various scales have been devised to measure either or both of these dimensions of social desirability. The earliest was the work of Edwards (1957) who felt that he was getting at fairly general cultural norms of social desirability with the Edwards SD Scale. His work was criticized by others who said that the SD scale was more of an acquiescence measure or a personality test.

Similar problems have plagued the various other scales which have been developed, such as those of Stricker (1963), Crowne and Marlowe's MC Scale (1964), and Schuessler *et al.'s* RD16 Scale (1978). Those that are supposed to measure social desirability as an item characteristic often measure facets of personality and *vice versa*. Even the validity of the MC Scale, widely used to measure an individual's need for approval and to identify those people most likely to distort responses in a socially desirable direction, has been questioned. MC scores were found to indicate actual socially desirable behaviour rather than the tendency to distortion of responses they were supposed to represent (Bradburn and Sudman, 1979).

A more promising approach was taken by Phillips and Clancy (1972) who conceived of social desirability as both an item and a personality characteristic. Hence, they saw a need

for respondents to rate the desirability of individual items (whereby ratings on desirability might vary from item to item), as well as a necessity to measure need for social approval (with a modified MC Scale). The latter would tap a personality characteristic independent of item desirability ratings. Phillips and Clancy found that trait desirability ratings were usually independent of need for approval scores. Thus, "respondents who have a high need for social approval do not rate the desirability of the various traits differently than do respondents with lower need for approval" (De Maio, 1984:263). De Maio cites further work by Gove and Geerken (1977) which substantiates Phillips and Clancy's findings.

Phillips and Clancy's study was important because people who were surveyed (by telephone, incidentally) and asked to rate themselves on such matters as happiness, religiosity, and so on, were also asked to rate these subjects on their social desirability. They also completed a need for approval measure. Their own answers to the questions were found to be more closely related to their social desirability rating of the subject than they were to their need for approval scores. In addition, they found that item characteristic desirability had a stronger effect on women than on men. Thus, this study has provided evidence that the item characteristic dimension of social desirability is a more important determinant of responses people make than is the personality characteristic dimension (deLamater, 1982).

De Maio (1984) has pointed out several other sources of confusion and inconsistency in this literature. In addition to the sloppy and inconsistent use of operational definitions of social desirability which include "acceptable" versus "desirable" behaviour, a need for social approval, and a tendency to respond desirably, she calls into question the reliability of the operational definitions of social desirability used in the various scales. She does so because respondents in the surveys may not have been sufficiently informed of what was expected of them. For instance, some judges of items were allowed to interject their own meaning for social desirability (e.g., the MC), while others were told to rate desirability independent of their own attitudes (e.g., RD16). Some judges may have confused social desirability with the concept of self-esteem (Phillips and Clancy, 1970). Others were given excessively vague coding instructions and were told to make an unnecessary distinction between coding for social

desirability and coding for threat (Sudman and Bradburn, 1974).

Because of findings that item characteristics are more important than personality characteristics as predictors of the social desirability of responses, De Maio proposes that further research be undertaken utilizing an "item-centred approach, hypothesizing that specific items will evoke a response consistent with social desirability as a function of their loading on social desirability . . ." (p. 264). This is the approach followed here. Further, as De Maio also recommends, this thesis attempts to deal with attributes of item desirability by proposing rules for judging "whether the content of individual items reflects positive, or good, consequences" (p. 268), and to determine whether items differ in the extent of their goodness or badness.

Other Factors in Social Desirability

The social desirability of answering a particular item in a particular way is not the only consideration in distortion of responses to enhance presentation of self. Although Sudman and Bradburn (1974) concluded that, in general, task variables such as social desirability were more important sources of response effects than respondent or interviewer characteristics, the latter were also influential in producing differential responses when social desirability was a factor. Interviewer characteristics were noted to interact with social desirability for attitudinal items, as in the case of larger response effects for white than for black interviewers. Weiss (1968) found social desirability bias to be associated with rapport with the interviewer on behavioural items, and with status similarity with the interviewer on attitudinal items. A later study of responses to questions about behaviours considered sensitive found that prior expectations of the interviewer underlay at least some of the social desirability bias in responses (Sudman *et al.*, 1977).

Respondent characteristics were reported to interact with social desirability in Phillips and Clancy's (1972) study which found the response effects greater for women than for men. But other researchers, such as Gove and Geerken (1977) in their examination of mental health measures, have not found evidence that social desirability bias is systematically related

to such independent variables as sex, race, education, occupation, or income (de Maio, 1984). Thus, the issue of how social desirability may interact with respondent characteristics is still unsettled. Sudman and Bradburn (1974) speculate that where there has been an interaction between sex or race and social desirability bias, it may have been due to an interaction between sex or race and the topic of the question, since the strong possibility of giving a socially desirable answer tends to arise on questions about racial, ethnic or sexual attitudes.

Other variables which have been found to interact with social desirability in affecting responses are the form of the question (closed versus open-ended) and saliency of the topic (Sudman and Bradburn, 1974). Saliency and threat (which is undoubtedly related to social desirability) were found to operate in opposite directions, i.e., "the higher the threat, the greater the response effect, but the higher the saliency, the lower the response effect" (Sudman and Bradburn, 1974:66). This latter finding of low response effects with high saliency is in line with Sudman *et al.'s* (1965) opinion that personal interviews could be as effective as self-admininstered questionnaires in obtaining answers undistorted by social desirability or self-aggrandizement biases if the respondent feels very strongly about the issue.

Most of these variables which have been found to interact with social desirability, as well as the effect of social desirability of the item itself, have not been systematically explored across types of interviewing situations. Nor will these interaction effects be addressed in detail in the following analysis. However, one potential interaction (between desirability and salience) is very relevant to predicting where mode effects might be expected, and is discussed with the hypotheses at the end of this chapter.

The Interviewer as a Factor

It is only logical to expect social desirability to have more influence in a face-to-face interview than in a telephone interview simply because there is another person present. The interviewer can be seen as an evaluator as well as an interviewer. But aside from the fact that this interviewer is likely to be providing nonverbal feedback cues to the respondent when he or she answers a question, it appears that there are certain qualities inherent in direct interaction

with another human being which may also be working to encourage more socially desirable responses. These may be seen to encourage answers which might be termed "accommodative" to the social situation of the interview.

That the interviewer is simply more important as a factor in the face-to-face interview than in the telephone interview would be deduced from research in social psychology which indicates that a stimulus is used "as a cue in making attributions" to the extent that it has "perceptual or emotional salience" (Sillars, 1982:82). Furthermore, Sillars cites research by Nisbett and Ross to explain that such salience is a matter of information exerting a disproportionate influence on social inference because it is "vivid", "emotionally interesting, concrete and imagery-provoking, and *proximate* [italics mine] in a sensory, temporal, or spatial way" (1982:82). The interviewer in a personal interview, by virtue of being there, of being "salient", has the respondent's attention focused on him or her, and there is a resultant tendency for the interviewer to be seen as "responsible" for surrounding events.

What lies behind such human attribution tendencies is unknown, but if such psychological mechanisms are really operating in face-to-face interactions, then it follows that what one says to an interviewer face-to-face may well be felt to be of greater weight interactionally than what one would say to a telephone interviewer. Hence, it becomes more of a necessity to give a socially desirable response to a person who is present.

There is also a seemingly unconscious process occurring in face-to-face interactions which can only be termed "accommodative" behaviour, illustrated by convergence of speech styles which has been observed on a number of descriptive levels and which cannot be explained by the normative demands of the situation (Giles and Smith, 1979). The usual aspects of speech examined in these linguistic accommodation studies are termed "noncontent" speech behaviours in that they concern how speech is produced, not what is said. Examples of the sort of features examined are pause and vocalization durations, speech rate, vocal pitch and intensity, and pronunciation (Street and Giles, 1982). Experimental manipulation of an interviewer's speech utterance length by doubling it, for example, produces a doubling of the interviewee's speech utterance length (Matarazzo and Wiens, 1972).

The explanation for such effects is unknown, but as Giles and Smith (1979) point out, one can only make the assumption that shifts in one's speech style to match that of the other have the purpose of encouraging further interaction and decreasing perceived discrepancies between actors. There is also the possibility of divergence as well as convergence in speech within dyads, and situations exhibiting shifts of this sort are likely to be those in which there is a role or power discrepancy between participants (Street and Giles, 1982).

How one views reasons for these changes is likely dependent on one's view of human nature. Giles (1980) has reviewed several explanations for convergence phenomena. They may represent a desire to increase similarity between people to increase attraction and intelligibility, and are, therefore, more or less standard operating procedures when strangers meet. Another explanation he mentions is that increasing similarity leads to greater perceived predictability. People may measure supportiveness of others by the degree of convergence behaviour they demonstrate.

What these kinds of accommodative behaviours might indicate for proposed differences between face-to-face and telephone responses is not entirely clear, but there certainly seems to be a mutual accommodation process taking place in face-to-face interactions of which we are seldom aware. Given this atmosphere of what is, in essence, cooperation, we might expect that other accommodative behaviours would be more likely in personal interviews, too. These could include trying to give more complete answers to questions, deciding to give an answer instead of responding "don't know", or, in other ways trying to please the interviewer.

Whether the same kind of linguistic accommodation would occur as readily in telephone interactions has not been examined as yet. It is known that we use voice to judge others on many traits. For example, voice alone can produce astonishingly high inter-rater agreement on judgments of personality, though these are not necessarily accurate. Voice can also produce fairly accurate judgments of age, sex, and even emotion. Less accurate judgments of physical appearance, height, complexion, vocation and political preference are made on the basis of voice (Brown and Bradshaw, 1985). It was found that people do use

voice to make judgments of social class and that class stereotypes produce resultant evaluative ratings on friendliness, trustworthiness, and likeableness (Sebastian and Ryan, 1985). In interviews, vocal characteristics have also been found to have important effects on response rates and facilitation of understanding of questions (Oksenberg and Cannell, 1988; Oksenberg *et al.*, 1986).

Some work by Brown, Giles, and Thakerar (cited in Brown and Bradshaw, 1985) provides indirect evidence, though, that a social process such as accommodation may not take place so easily from voice alone. They replicated previous findings that different people make the same judgments of others' competence and benevolence based on their speech rates. However, they found that these ratings changed considerably when a context was provided. Thus, if subjects were given a reason for slow speech rates (usually judged as a sign of decreased competence and benevolence), slow speech was then judged as showing increased benevolence and competence.

It is likely that this experiment illustrates some of the differences between face-to-face and telephone-mediated interaction situations. Evaluations of the other can be very different if nonverbal (visual) cues are present because these cues would help to provide a "context" for what the voice says and how it says it. Without nonverbal communication, or perhaps simply without being able to see the other, evaluative reactions to voice alone would tend to be fairly stereotyped.

In attempting to assess what difference it might make if a "sensitive" question is asked over a telephone, we would need to know to what extent people routinely evaluate the owner of the voice on the other end. It is known that evaluations <u>can</u> be made solely on the basis of voice, and that there is a surprising uniformity in the content of these evaluations. If such evaluations of the speaker are normally part of a telephone interaction, we might expect a certain reduction in variability of responses, particularly if there is a stereotyping in evaluations. But if telephone voices do not routinely arouse a need to visualize and assess the speaker, we would expect greater variability in responses. Even with some fairly concrete image of the speaker based on voice alone, though, we could not expect such an image to have

as strong an impact on a respondent's answer as a live interviewer providing nonverbal cues.

D. Role of Nonverbal Cues

The interviewer, simply by being present, has a salience to the respondent, as already mentioned, and thereby exerts more influence on social inference than would occur without that physical presence. This saliency of the other, the interviewer, is no doubt bolstered by the nonverbal feedback cues which he or she is constantly supplying to the respondent, whether consciously or not. The importance of this nonverbal information in interaction is so great that Goffman has even stated that reliance on words alone "can conceal the interactional facts" (1981:36). Non-verbal communication, also referred to as "metacommunication", provides the cues which contextualize a message and indicate how it is to be taken. It does so by providing the means through which relationships "are defined on such dimensions as power and affect" (Sillars, 1982:92).

The kinds of communication which occur nonverbally are several, but nonverbal signals all serve either or both of Birdwhistell's (1970) "informational" or "integrational" functions. The former is concerned with the passage of information. The latter "keeps the system in operation, regulates the interaction process, cross-references particular messages to comprehensibility in a particular context and relates the particular context to the larger contexts of which the interaction is but a special situation" (Short *et al.*, 1976:44). However, verbal cues could also be said to serve these same two functions of information and integration, only in different ways.

Nonverbal signs are best suited to the transmission of subtle differences, for example in intensity or extensiveness, because the information they convey is probabilistic rather than invariant, and continuous rather than discrete (Scherer, 1980). This means that nonverbal signs have meanings which are not entirely certain, and that they can communicate gradual differences in the size and quality of what they are referring to. Thus, they are better for information one wishes to "try out" on a receiver, "keeping open the option to disclaim the intention to communicate this information" (Scherer, 1980:227).

On the other hand, information one is certain of may be better communicated via the invariantly coded verbal signs. Just as a person would make use of nonverbal signs to communicate such uncertain information, he or she would use such nonverbal feedback signs on the receiver's part to see how the uncertain information was being interpreted and evaluated. Thus, nonverbal cues should be expected to play a greater role in situations of uncertainty, where words alone do not convey all the information needed in a situation.

The specific ways that visual cues function in interaction to accomplish the overall informational and integrational goals have been discussed by Reid (1977) who is summarizing work by Argyle (1969). These functions include those of showing mutual attention and responsiveness or evidence that the other person is attending to what you say, and which might include such signs as eye contact or head nodding. Such signals may have an important role in encouraging accommodative or affiliative behaviours in face-to-face interaction. One experiment demonstrated that head nodding was able to dramatically increase the length of interviewee's speech (Matarazzo *et al.*, 1965). But whether it is simply attention which is conveyed and which the interviewee finds rewarding, or whether approval is the information conveyed is not clear. Thus, cues functioning as attention demonstrators might just as easily be included with interpersonal attitude communication functions in terms of what they accomplish.

The next function of visual cues is that of channel control, or indicating the way participants should take turns speaking and listening. Channel control is usually communicated by gaze or body shifts (Scherer, 1980). Absence of these channel control cues may very well be responsible for the the often-noted shorter responses to open-ended questions in telephone surveys. Interpersonal attitude communication (to show attitudes or intentions) is another function. The face, gestures, body movement, posture, distance, and eye contact can all convey personality or emotional arousal, while status can be transmitted clearly by posture (Scherer, 1980). The illustration function allows people to accompany and illustrate what is said, primarily by gestures. Finally, the feedback function shows if the other understands, believes, is surprised, agrees or disagrees, is pleased or annoyed with what is

said.

This latter feedback function appears to be a catch-all and really includes the others, except that because feedback cues are reactive, they are being considered from the point of view of the person spoken to, not the speaker. These feedback cues are the ones that should be most responsible for producing differences in response patterns between face-to-face and telephone interviews. As previously mentioned, these feedback cues provide signals of attention, comprehension, and evaluation (Scherer, 1980) and are being mutually provided to each other by the interviewer and the respondent in the interview. If the respondent sees a smile or nod, he or she is encouraged to pursue that line of thought or to provide more information. It has been demonstrated, for example, that warmth created by a handshake or smiles during an interview can lead to more disclosure of intimate information (Siegman, 1980).

Or, the respondent can use that "special way of knotting up the face to convey the fact that [he does] not understand what it is a speaker seems to be trying to convey, and that a rerun is in order" (Goffman, 1981:36). The interviewer, thus, receives a signal that clarification is needed, or in a similar way, the respondent can learn that his answer has not been understood, possibilities which are less likely to occur in telephone interviews and which would probably lead to different responses. In the same way, feedback in terms of the evaluative cues given by the interviewer can convey signs of agreement or disagreement, liking, e.g., by leaning forward with the torso (Scherer, 1980), or disapproval (the proverbial raised eyebrow), doubt (shaking one's head or shrugging), all signs which could lead face-to-face respondents to give a different answer than their telephone counterparts.

Although most of the above feedback cues are thought of as ways which might increase response, both in amount and probably quality, it is conceivable that the greater amount of information conveyed face-to-face, its greater channel capacity, might also act to inhibit disclosure. If the respondent is picking up disapproval or untrustworthiness cues visually, or perhaps can visually perceive too great a status differential between him or herself and the interviewer, even less information may be shared in personal interviews than in

telephone interviews.

There is also some evidence that there may be limits to the effect that all the "encouraging" kinds of nonverbal feedback signals can have on increasing response. The experiment which manipulated the effect of interviewer warmth on response, for example, showed that there was no across-the-board facilitating effect on interviewee communication as a reinforcement-based attraction model would have predicted. Findings that warmth on the part of male interviewers actually led to a decrease in female interviewee productivity were felt to support the equilibrium model of interpersonal intimacy proposed by Argyle and Dean (1965). Thus, "participants in social interactions tend to establish an appropriate level of intimacy by means of a delicate balancing or equilibration of various affiliative behaviors such as eye contact, physical proximity, warmth, topical intimacy and so forth" (Siegman, 1980:89-90).

There is certainly evidence that nonverbal signs should play a more complex role in communication during interviews than is normally recognized. Although they can function in an almost infinite number of ways, it is likely that they are most useful and influential in particular types of situations because of the unique types of information they can convey. With respect to the subject of this thesis, responses to sensitive questions, nonverbal signs which are absent in telephone interviews should be important in affecting responses to questions that respondents are unsure how to answer. Unsureness could derive from either uncertainty about which answer is most morally correct, or from simple comprehension difficulties. Where there is certainty, nonverbal signs would, in effect, simply be transmitting redundant information. Thus, they are not necessary to respondents in situations of certainty, and we would not, then, expect mode differences in responses. Non-verbal signs could also contribute to mode differences because they can convey approval or provide some type of reinforcement which encourages greater response.

E. Communications Medium Experiments

In order to make predictions about specific types of differences expected in telephone and face-to-face interview responses to items prone to social desirability, it might be useful to review some of the work done on effects of medium on communication. About thirty fairly rigorous laboratory experiments have been conducted which explore how communication is affected by use of telecommunications media in various types of situations of cooperation and conflict tasks (Williams, 1977). Unfortunately, these experiments were not directed to comparisons of telephone and face-to-face responses to survey questions *per se*, so any results can only be assumed to apply to interviewing situations.

For cooperative tasks involving problems with objective solutions, there was little difference in performance between telephone and face-to-face conditions. Various tasks involving conflict, however, did show medium effects in many ways. These conflict tasks involved such varied situations as simulated labour/management negotiations, arguing a case which did or did not correspond to the subject's own opinion, decisions to use competitive or cooperative strategies in mixed-motive games, how confederates were evaluated in Prisoner's Dilemma games, and opinion change when there was initial conflict of opinion.

For these tasks, the medium was felt to affect results because the exact relationship between the participants is important in conflict situations, whereas in cooperation or information conveyance tasks, the relationship is not so critical to the outcome. General findings were that audio-only conditions produced more opinion change and made the strength of an argument the primary determinant of outcome in negotiations. On the other hand, face-to-face conditions produced more use of cooperative strategies in games and emphasized affective content in messages. Thus, under audio-only conditions, the content of the message is all that appears to matter. Seeing the other person, though, seems to make people take that person into account to a greater extent and emphasizes the emotional content in messages. These results are reminiscent of those in the experiment of Brown, Giles, and Thakerar cited earlier in which evaluations of the other's competence and benevolence based on speech rate changed considerably if a context was provided (cited in Brown and Bradshaw,

1985).

The studies of interpersonal perception show that communications medium can affect peoples' impressions, and that media with more nonverbal cues produce more favourable impressions, (Williams, 1975), but that the effect is not very strong (Williams, 1977). Williams (1977) concludes that there is no unitary theoretical explanation for the differences that have been demonstrated, since those which emerged could not consistently be explained either by the role of nonverbal cues or by efficiency of information transmission. One explanation he proposes is that nonverbal cues may be conveying information which is already being communicated in other ways. Thus, when nonverbal cues are redundant in this sense, they would not have the expected effects. He feels that the only sort of explanation which can encompass all the findings is one which sees the whole nature of communication itself as being altered by the absence of some channels. Thus, "while face-to-face we see others as real social beings, with individual personalities, wishes, feelings, and aspirations," but over the telephone "we treat others more like semimechanical objects which can be ignored, insulted, exploited, or hurt with relative impunity" (Williams, 1977:972).

Similar ideas were explored in depth by Short, Williams, and Christie (1976) who proposed that each medium of communication has its own level of a quality they call "Social Presence", a hypothetical construct which means the "degree of salience of the other person in the interaction and the consequent salience of the interpersonal relationships" (p. 65). Basing their line of thinking on Argyle and Dean's (1965) equilibrium model of interpersonal intimacy, Short *et al.* propose that Social Presence of the communications medium should be considered as another factor contributing to intimacy, and, like the other factors, it is juggled to create the right level of intimacy. Because different tasks require different degrees of Social Presence there is no way to describe any overall general effect of communications medium. The suitability of any medium for a particular type of interaction will depend upon both the degree of Social Presence of the medium and the degree of Social Presence required by the task. They predict that the greatest distortions compared to face-to-face communication would occur when a medium with a low degree of Social Presence was used for a kind of

interaction requiring a high degree of Social Presence.

F. Hypotheses

It follows that giving an honest answer to questions which have socially desirable or undesirable answers is a task which is subject to the influence of Social Presence and which should, thus, be affected by the medium of communication used. There is something to be gained or lost by answering desirably or undesirably in a personal interview, that is, the esteem of the other. On the other hand, more trust should be present face-to-face, while the telephone provides anonymity, factors encouraging potential disclosure of undesirable information.

There is not much here in the way of guidance for predicting exact mode effects on responses because rules have not been deduced for maintaining an "equilibrium of intimacy," nor for gauging the relative strength of various types of nonverbal cues on interaction. There is also the possibility that if communication using different media is truly different in its very nature, then interaction under these circumstances also differs in kind and is, therefore, subject to potentially distinct ways of maintaining an equilibrium of intimacy. However, from impressions gained in the process of reviewing the foregoing material, several testable hypotheses have emerged about where mode of interviewing is likely to affect responses to questions which have socially desirable or undesirable answers.

The following hypotheses, which are tested in the next chapter, are derived from two main principles, 1) that the extent to which communications medium will affect responses to items is a function of the degree of social desirability involved in giving particular answers, and 2) that the medium is more likely to affect answers to items which ask the respondent to do a little extra work, for example, when the respondent must search his memory in order to answer. The ultimate reason for the existence of an effect of communications medium has to do with the peculiar nature of nonverbal signs. The hypotheses are:

1. The degree of social desirability involved in answering a question in a particular way determines whether mode affects response.
   a. For items having a very socially desirable or very socially undesirable answer, there will be no difference between modes in social desirability of the response.
   b. For items which have a somewhat socially desirable or a somewhat socially undesirable response, face-to-face respondents will give more socially desirable responses.
   c. For items which do not appear to have any social desirability component, there will not be a difference between face-to-face and telephone responses.
2. The actual physical presence of an interviewer in personal interviews will have a positive effect on the respondent's motivation to answer questions.
   a. There will be less nonresponse on all items in personal interviews.
   b. There will be differences between modes on means or proportions for items of neutral social desirability if recall is required.
3. Mode effects based on both degree of social desirability of the item and the respondent's motivation level will be reflected in variability of response.
   a. There will be no difference between modes in variability of response for very socially desirable or very socially undesirable items.
   b. There will be more variability for telephone respondents on somewhat socially desirable and somewhat socially undesirable items.
   c. For items without a social desirability component, variability will not differ between modes.
   d. There will be more variability in personal interview items of neutral desirability which require recall by the respondent.

Table 1 summarizes the predictions for mode effects which derive from these hypotheses.

On very socially desirable and very socially undesirable items it is expected that there will not be mode-related differences in responses because there will be a great deal of ego-involvement for the respondent in giving the socially desirable response. Taking such a

Table 1

Predictions

Category	Mean or Proportion	Variability	Nonresponse
1. Very Socially Desirable	Same in both modes	Same in both modes	Significantly more for telephone
2. Somewhat Socially Desirable	Significantly in more socially desirable direction for face-to-face	Significantly more for telephone	Significantly more for telephone
3. Neutral Desirability	Same but significantly different if recall required	Same but significantly more for face-to-face if recall required	Significantly more for telephone
4. Somewhat Socially Undesirable	Significantly in more socially desirable direction for face-to-face	Significantly more for telephone	Significantly more for telephone
5. Very Socially Undesirable	Same in both modes	Same in both modes	Significantly more for telephone

position does not imply that there will be no distortion on the part of respondents, but that such distortion should not be any different over the telephone than it is in person. This type of item reflects strong social norms. Most people probably do behave or think in the socially desirable way for these items, but even if they do not, it would be the strength of the norm, its saliency, which provides the impulse towards a conforming response.

The presence of the interviewer and the associated nonverbal communication, then, have relatively less effect on responses to items with very desirable or undesirable answers. This position is consistent with previous findings of lower response effects for highly salient items, as well as with the relatively stronger impact of the message contents for audio-only conditions in the communications research. Compared to the item itself, the Other, or the interviewer, is having less of an effect on response for such very desirable or undesirable items. For these items, information which would be provided by the interviewer's nonverbal evaluation cues would be redundant to information the respondent already has. Therefore, the nonverbal cues would not be so necessary to the respondent in formulating his or her response and would be ignored. Such a position is also consistent with findings by commitment theorists that involvement by subjects decreases the ability of a source to change attitudes (cited in McGuire, 1985).

Because of the strength of the social norms behind these very desirable or undesirable responses, there would not be any statistically significant differences in variability by mode of interviewing. Under both conditions, responses should be converging on the socially desirable option, although there may be a few pranksters in the telephone sample who take advantage of the telephone's anonymity and traditions of playing tricks on the phone to play jokes on the interviewer. Such contributions of "silly" responses should be minimal, though.

Possibly, the "mechanics" of communicating the response categories for a question over the telephone can contribute to slightly more variability for telephone responses which is unrelated to desirability content. Thus, the need to repeat the endpoints of a scale on the phone instead of being able to show a card with printed response values, may increase the saliency of scale endpoints and consequent variability in response.

It is on the somewhat socially desirable and undesirable items where we should expect response differences between the face-to-face and telephone samples. Respondents in the personal interview situation should give socially desirable responses to these items significantly more often than do telephone respondents. The evaluative nonverbal feedback cues provided by the interviewer under these conditions where the norm is not so strong as to "force" the socially desirable response, or where the respondent is unsure of the socially desirable response, would be responsible for the mode difference. Non-verbal cues, then, would be conveying a needed "context" which is not redundant information to the respondent.

In addition, attention-showing feedback cues are rewarding in themselves and encourage the respondent to please the interviewer by selecting the more socially desirable response. These predictions are consistent with the ideas on Social Presence of Short *et al.* (1976) in which medium most affects tasks requiring assessment of the other's reactions.

Variability in face-to-face interviews should be significantly lower for these items than it is for telephone interviews, again because there is more tendency to converge on the socially desirable response option. Telephone respondents do not have the nonverbal information to help them choose the desirable option, nor do they have the nonverbal motivation cues to encourage this choice.

For items without any social desirability connotation, there should not be any *a priori* reason to expect differences between face-to-face and telephone responses. Thus, means should not differ on scale items between samples, nor should proportions choosing either response alternative on nominally scaled questions, unless there is some factor such as complexity of the question operating to cause mode differences. Because of the more rewarding social atmosphere of the face-to-face interview with its wealth of nonverbal feedback, there should be more of an effort on the part of the respondents, though, to recall details of behaviour when queried about these. A byproduct of such increased effort should be more variability of responses face-to-face, which would be in line with previous findings of greater variety and quantity of response face-to-face on open-ended items.

Finally, there would be significantly more non-response in telephone interviews on all types of items simply because of the inherently social nature of the personal interview with its innately rewarding interaction. It is also more difficult not to give some answer in a personal interview because of "demand characteristics" of the situation with the compelling physical presence of the interviewer being more important in this regard than the norm of telephone behaviour entailing interactional obligations (Frey, 1983).[8]

[8]The importance of this motivating effect of the interviewer was documented in an experiment which compared formal and informal interviewing styles. Motivational level turned out to be a better explanatory principle for the effect of interviewing style on responses than was tendency to ingratiate onself when items had no social desirability component (Dijkstra, 1987).

III. Data and Methods

A. Data Set

The data used in this study come from the 1988 All Alberta Study conducted by the Population Research Laboratory (PRL) of the University of Alberta's Sociology Department. This study was the second province-wide survey undertaken by the PRL. Their previous surveys had been limited to face-to-face interviews in Edmonton only, the Edmonton Area Study (EAS). The 1988 survey was the first, though, to include both a face-to-face and a telephone sample in Edmonton itself, which made it a useful body of data to examine for mode effects.

The Edmonton face-to-face portion of the survey was a random sample of the population of all persons 18 years or older living in a dwelling unit which had been enumerated in the 1987 Civic Census. The sampling frame was a computerized list of addresses from the 1987 enumeration, from which a total of 620 addresses was randomly selected. For each household, one eligible person was selected as the respondent for the one hour interview, with interviewers attempting to obtain equal numbers of male and female respondents. The final sample size for the Edmonton face-to-face group was 464, making it comparable to that of previous Edmonton Area Surveys. Inappropriate addresses (e.g., temporary residences, nursing homes) were deleted. With 584 households found usable, the response rate was calculated at 79.5% (Kinzel and Adebayo, 1988).

The telephone sample in Edmonton consisted of 116 individuals selected by random digit dialing from an original sample selection of 304 telephone numbers. The three digit prefixes of telephone exchanges in Edmonton were assembled to generate seven digit telephone numbers. Then there was a random selection of exchange and selection from all the five digit working banks in the reverse directory accomplished by fixing a random two digit number to selected working banks. Businesses were eliminated by checking numbers in the reverse

directory. With 161 usable numbers, the response rate was calculated at 72%.[9] As in the face-to-face portion of the survey, attempts were made to obtain equal numbers of male and female respondents.

The two samples were compared on the demographic and economic variables, age, sex, marital status, years of schooling, and household income. It was found that the percentage distributions of respondents in various categories of these variables did not differ significantly by mode when a chi-square test was used (Lalu, 1988). I also tested (using t-tests) for differences between samples on means for age, years of schooling, and income and did not find any significant differences.[10] Thus, any differences in responses between the telephone and face-to-face respondents should not be a function of differences in the composition of the samples.

The telephone questions were all taken from the face-to-face questionnaire, but the length of the telephone questionnaire was shorter (by about half) with some questions being modified slightly for administration over a telephone. There were, however, no modifications to response categories or scales. The particular questions chosen for analysis in this thesis were ones for which there were no modifications for telephone administration.

Unfortunately, the ordering of the questions in both surveys was not exactly the same. Because the telephone questionnaire was shorter, certain topics were omitted, with the result that different subject matter areas were juxtaposed in the two questionnaires. Where there were many questions asked about a topic, for example, ACCESS television, question order was identical, though.

A wide range of topics was covered which reflected interests of researchers at the University of Alberta, interests of outside agencies participating in the survey, replications of

---

[9]From a discussion with Cliff Kinzel, PRL Research Technologist, on September 28, 1989.

[10]When I retested for similarities in percentage distributions of respondents on background variables using a chi-square test, I obtained p values of .99 for age, .95 for sex, .91 for marital status, .29 for years of schooling, and .48 for household income. Results of t-tests showed p values of .71 for age, .79 for years of schooling, and .29 for household income.

questions from previous years, and the special topic for the 1988 survey which was "coping with the recession." Subjects covered in both modes included demographic characteristics, gender roles, the Alberta economy, attitudes about post-secondary education, free trade, discrimination on the basis of sexual orientation, television viewing, road safety, and personal travel.

Face-to-face interviewing was carried out by thirty-eight interviewers (five of whom were male) starting in the first week of February, 1988 and continuing for six weeks. Telephone interviews were primarily conducted by four female interviewers, three of whom had also worked on the face-to-face interviews. All received separate training for the telephone work. Two other women, one of whom was a face-to-face interviewer, helped with follow-ups on refusals and cases involving language problems.

B. Choice of Variables

The best way to test the previous ideas about how the mode of interviewing affects responses would be to focus on questionnaire items which vary in the social desirability of particular response patterns. There would have to be some criteria, however, for choosing questions from the survey which were likely to be subject to greater and lesser amounts of bias due to social desirability. Ideally, there would be a range, from items to which a certain answer would be considered very socially desirable to items for which a certain answer would be thought of as very socially undesirable. Differences betweeen the telephone and face-to-face respondents' answers to these items could then be sought.

The problem lay in how to choose the questionnaire items which had socially desirable or undesirable answers. What is defined as desirable can vary between individuals, between populations or even segments of a population. Desirability can also vary from one time period to another within the same population. The most valid assessment of social desirability, for the present purposes, would have to be one based on current norms prevalent in Edmonton, where the actual survey had been carried out.

To determine these norms, it was decided to conduct a small informal survey to assess the social desirability of giving particular types of responses to particular questions from the 1988 survey.[11] A panel of eight experts, individuals who have conducted their own survey research in Edmonton or who have had experience analyzing survey data from Edmonton, as well as twenty-six nonexperts, living in Edmonton, comprised the sample for this informal poll. The experts were faculty members of the University of Alberta Sociology Department, the Faculty of Business, and data analysts from the PRL and the University of Alberta Department of Computing Services.

The non-expert group was made up of ten of the interviewers who had worked on the survey[12] and sixteen friends and neighbours of the author (of both sexes and ranging in age from the early twenties to about eighty). There was a definite middle-class bias to the sample, which could make some difference to an assessment of social desirability if the majority opinions in the society were not the same as those of the middle-class respondents. However, since the panel of experts had experience surveying the range of types of people in Edmonton, it was felt that their opinions could encompass views which were being neglected.

Participants in the informal survey were given a list of twenty-eight questions, and possible answers, taken from the 1988 EAS. This list of questions with instructions to the respondents can be found in Appendix One. Participants were asked to rate the answer to each question on a scale of 1 to 5 which indicated how socially desirable they felt that particular response was to that particular question. A "1" meant "very socially desirable", "2" "fairly socially desirable", "3" "social desirability would not be a factor in how a person answers this question", "4" "somewhat socially undesirable", "5" "very socially undesirable." In addition, the experts were provided another rating category, "9", for situations where they felt the answer was socially desirable for some people, but undesirable for others.

---

[11]I am indebted to Dr. M. Gillespie for suggesting this strategy.

[12]Surveying interviewers represented a compromise solution to a problem raised by one expert (a non-participant) who felt that interviewers should be administered the same questionnaire as respondents in the EAS. This would be done to determine whether respondents in face-to-face interviews were truly matching their responses more to what interviewers saw as "correct" answers.

The basis for choosing items for the informal survey was simply my own feelings about which questions and answers might exemplify attitudes and experiences which people in Edmonton would agree were desirable, undesirable, or neutral in desirability. If there was a choice between two questions to include, and if there had been an indication of mode differences for response patterns on one of them in a preliminary report (Lalu, 1988), then that question with differences was chosen.

The subject matter of most EAS items appearing on both the telephone and personal interview questionnaires was not, for the most part, inherently undesirable. Some of the most obvious questions for measuring undesirable behaviours, such as those on smoking and alcohol consumption, were not asked in both modes. Thus, the choice of obviously undesirable items was somewhat limited.

An alternative strategy could have involved asking participants to rate the social desirability of an "undesirable" response to a question about something obviously desirable, e.g., "How often do you get together with friends", (Response = Never) in order to present a greater number of undesirable choices. However, such a strategy could be very confusing to the rater, and would not ensure certainty in interpretation of ratings. A "4", for example, might mean that the rater thought either that the response was undesirable or perhaps that the idea of getting together with friends was undesirable. Thus, to minimize confusion, responses to be rated were kept consistently positive (e.g., yes, often).

Even with these precautions, there was still confusion among the non-experts when they were rating the responses. Despite the fact that I was present while friends and neighbours filled out their questionnaires and could answer their questions, it was obvious that some of them occasionally answered in a way that was opposite to their intentions. The problem was that people forgot that they should be rating the response to the item, and not the subject matter of the item itself. When people realized they were confused, we could clarify what they were supposed to be evaluating, but the problem still went uncorrected on a few questionnaires.

The same problem occurred in the interviewers' questionnaires despite, or perhaps because of, their familiarity with the questions. Thus, interviewers would forget that they were supposed to provide their own, personal rating of the response and would rate according to how people they had interviewed would assess that response, or the subject matter of the question. Difficulties which interviewers experienced were probably exacerbated by the fact that in many cases I could not be present while they completed their forms. As in the questionnaires filled out by friends and neighbours, there were obviously some ratings made which were not truly those intended.[13]

Despite the difficulties, there were a number of question/response items which did elicit agreement among raters. The questions which eventually were used in further analyses were those for which there was most agreement among experts and non-experts and where the format of the question and response categories were the same in both modes so as to avoid effects which might be due to differing question formats. In all cases, questions chosen had a clear majority of judges overall rating them similarly. But it was not always the case that a majority within each category of raters, (experts, interviewers, and others), had assigned the same social desirability value to the item. The actual tabulations for how the questions and responses were rated for social desirability can be found in Appendix Two.

Variables chosen for analysis, classified by degree of social desirability or undesirability, were:

Very Socially Desirable (rating number 1)

a. There should be stricter penalties for parents who do not properly secure their children in car safety seats. (Var061)[Variable label as it appears in the PRL's SPSSX systems file]

(Rated Response=Agree)

---

[13]The fact that there were a fair number of problems encountered in filling out the social desirability assessment questionnaire indicates that: a) the task is probably a more complex one for people than had been expected; b) more care should have been taken to ensure that questions about the questionnaire could be answered on the spot, or even that reminders about how to answer the questions should be provided intermittently (even though respondents appear to be coping adequately); c) the question contents are perhaps too ambiguous for many people for easy decisions about social desirability to be made.

b. In the past three years (or as long as you have lived in Alberta, if less than three years), have you or any other members of your household made any contributions to a Food Bank? (Var318)

(Rated Response = Yes)

Somewhat Socially Desirable (rating number 2)

a. It would be better for a pre-school child's emotional development if the mother didn't work outside the home. (Var273)

(Rated Reponse = Agree)

b. All in all how satisfied with life are you these days? (Var216)

(Rated Response = Satisfied)

Desirability Not a Factor (rating number 3)

a. I don't think the Federal government cares much about what people like me think. (Var336)

(Rated Response = Agree)

b. Are you aware of the television broadcast service? [provided by ACCESS Network] (Var165)

(Rated Response = Yes)

c. About how many hours in the past week have you personally watched ACCESS Network Television on cable channel 9? (Var172)

(Rated Response = Many Hours)

Somewhat Socially Undesirable (rating number 4)

a. There is a good chance that someone in my household will be unemployed in the next year. (Var339)

(Rated Response = Agree)

b. In the past three years (or as long as you have lived in Alberta if less than three years), have you or any other members of your household received financial help from relatives? (Var326)

(Rated Response = Yes)

c. About how many hours in the past week have you personally watched television? (Var169)

(Rated Response = Many Hours)

Very Socially Undesirable (rating number 5)

a. Punishments for impaired drivers are becoming too severe. (Var064)

(Rated Response = Agree)

b. In the past three years (or as long as you have lived in Alberta, if less than three years), have you or any other members of your household received Welfare assistance of any kind? (Var325)

(Rated Response = Yes)

c. It's okay to drive soon after having two or three alcoholic drinks (within 1 - 2 hours). (Var063)

(Rated Response = Agree)

C. Statistical Tests

For each of the variables listed in the preceding section, response patterns were compared for the face-to-face and telephone samples in several basic ways. Three types of fundamental differences were predicted. These were: 1. a more socially desirable response in face-to-face interviews, indicated by higher means (on response scales where higher values lie at the "socially desirable" end of the scale) or a larger proportion (of the more socially desirable response for dichotomous variables); 2. evidence of greater variability in telephone responses indicating less convergence on the socially desirable answer, but also in personal interviews for recall items without a desirability component (indicating accommodative and more motivated response); and 3. greater amounts of nonresponse among telephone respondents, indicating relative lack of motivation.

T-tests were used to determine whether the means on ratio or interval-level variables differed. This test was used for the hours of ACCESS TV and the hours of television variables. T-tests were also used to compare means on the seven-point scale items where respondents were asked to rate their agreement or disagreement with an item, thereby treating

these variables as interval level data. Variables looked at in this way included childrens' car safety seats, mothers of pre-school children working, satisfaction with life, Federal government caring, chance of unemployment in household, impaired driver punishments, and driving after drinking. A mode effect was considered to exist if the F ratio was significant at the .05 level.

For dichotomous response variables, proportions of respondents giving the more socially desirable response in both samples were compared using Fisher's exact test for the chi square test of independence.[14] If a significance level of .05 (one-tailed test) was obtained, this was considered evidence of a possible mode effect. The variables compared in this manner were contributing to a Food Bank, awareness of ACCESS TV,[15] receiving financial help from relatives, and receiving Welfare.

In order to determine whether telephone respondents showed less tendency to converge on the socially desirable response, the variability of response patterns was compared for telephone and face-to-face respondents on each of the variables. The test for this comparison of variability was an F-test of differences between variances, with a significance level of .05 or less (one-tailed test) considered to indicate more variability.

The predicted pattern of a higher amount of non-response for the telephone sample was checked for every variable. In each case, proportions of "don't knows" and "no response" were combined and compared to the proportion who did respond, again using Fisher's exact test and an .05 level of significance.

---

[14]Since not answering the question is considered to be socially undesirable, the proportion not responding, and proportion of "don't knows", were added to the proportion giving the undesirable response for purposes of comparison. These proportions of nonresponse were quite small in every case.

[15]The exact wording of this variable is, "The ACCESS NETWORK (an Alberta Crown Corporation) is involved in various activities. Could you tell me which ones you are aware of?" [respondent either mentioned ACCESS TV service or not]. A shortened version appeared on the informal survey.

IV. Results

It was predicted that interviewing mode would only affect the responses to items having somewhat socially desirable or undesirable answers, since respondents would rely more on the nonverbal evaluative cues of the interviewer. For very desirable or undesirable items, these cues are not needed because the respondent feels the full force of the social norms guiding his or her answer. Thus, we expect no difference by mode for very desirable or very undesirable items since nonverbal cues would only be supplying redundant information for these. We do, however, expect face-to-face means or proportions to differ significantly from telephone means or proportions, and to do so in the socially desirable direction, when the item has a somewhat desirable or undesirable answer. Furthermore, as a result of these same factors, we expect more variability among telephone respondents only on the somewhat desirable or undesirable items.

It was also predicted that nonverbal cues demonstrating attentiveness and even sub-conscious "accommodative" factors inherent in interaction would motivate respondents to answer in face-to-face settings. The result would be significantly more nonresponse on all items for the telephone sample. These same factors should also produce significantly more variability for face-to-face responses on recall items devoid of social desirability. This is because they also motivate effort to produce accurate and detailed response and, hence, greater variety in answers.

A. Preliminary Findings - Item Desirability Hypotheses

Analysis of mode differences for means and proportions show predicted results are obtained on nine of the fourteen items. That is, the initial findings on nine of the fourteen variables support the hypotheses (Table 2).

Very Socially Desirable and Very Socially Undesirable

The two variables which are considered to have very socially desirable responses, "car seats for children" and "contributed to Food Bank", do not show significant differences.

Table 2

Sample Differences on Means or Proportions

Variable	Telephone N=116	Face-to-face N=464	t-Value	p Value*
Rating #1 - Very Socially Desirable				
1. Car seats for children (1=Disagree, 7=Agree)	5.67	5.70	0.13	.45
2. Contributed to Food Bank last 3 years? - "Yes"	63.8%	65.3%		.84
Rating #2 - Somewhat Socially Desirable				
1. Better for pre-school child if mother doesn't work (1=Disagree, 7=Agree)	5.12	4.99	.68	.25
2. How satisfied with life are you? (1=Very dissatisfied, 7=Very satisfied)	5.44	5.72	1.99**	.02
Rating #3 - Neutral Desirability				
1. Federal government doesn't care (1=Disagree, 7=Agree)	5.17	4.94	1.27	.10
2. Hours watched ACCESS TV past week	.90	1.37	2.00**	.02
3. Aware of ACCESS TV - "Yes"	71.6	76.7%		.30
Rating #4 - Somewhat Socially Undesirable				
1. Received financial help from relatives past three years - "No"	87.1%	78.2%		.05
2. Hours watched TV past week	12.96	15.16	1.67**	.05
3. Chance someone in household unemployed next year (1=Disagree, 7=Agree)	3.35	2.78	2.21**	.01
4. Been unemployed past 3 years - "No"	70.7%	75.2%		.38
Rating #5 - Very Socially Undesirable				
1. Impaired driver punishments too severe (1=Disagree, 7=Agree)	2.27	1.87	2.34**	.01
2. It's okay to drive after 2-3 drinks (1=Disagree, 7=Agree)	2.29	2.25	0.21	.42
3. Received welfare past 3 years? - "No"	85.3%	89.9%		.22

* p values here are all one-tailed tests.

** Denotes significance at .05 level, one-tailed test.

although for both, the personal interview sample reflected greater social desirability than did the telephone sample. But face-to-face responses were not significantly different from telephone responses in terms of average agreement that there should be stricter penalties for parents who do not secure their children in car safety seats, or in terms of percentages who claim to have contributed to a Food Bank within the last three years.

Thus, these findings support the hypotheses for very socially desirable items. For such items, the mode of interviewing appears to be of little importance in influencing one's answer. People may pay more attention to the message than to the messenger (or to the question rather than to the interviewer) when there is a strong societal norm concerning the desirability of responding in a particular manner.

Two of the three very socially undesirable variables do not show significant differences between the samples, as hypothesized. However, the question asking if the respondent agrees or disagrees that punishments for impaired drivers are becoming too severe displays a significantly higher mean for telephone respondents (i.e., more agreement with the statement, indicating a more socially undesirable response). However, the telephone group does not differ from the face-to-face sample on the other drinking and driving question, "It's okay to drive soon after having 2-3 alcoholic drinks (within 1-2 hours)". Such a different pattern of responses on questions about similar subject matter may mean that the question wording or format may have played a role in the results. Additional analysis of the "punishments for impaired drivers" item is reported in the next chapter.

A pattern similar to the "okay to drive after 2-3 drinks" item, (a slight, but not statistically significant difference towards the "undesirable" for telephone respondents) was found for the question asking whether the respondent, or anyone in the respondent's household, had received welfare in the past three years. Overall, then, hypotheses about a lack of a mode effect on means or proportions for very socially desirable or very socially undesirable items are supported, except for the one item which asks whether the respondent agrees that punishments for impaired drivers are becoming too severe.

Somewhat Socially Desirable or Undesirable

In the "somewhat socially desirable" category on the other hand, there is more evidence of mode-related differences between samples. For example, one variable, "All in all how satisfied with life are you these days?", does show a significant difference ($t=1.99$, $p=.02$, one-tailed test) in means between the face-to-face and telephone groups in the direction predicted, with face-to-face respondents giving the more socially desirable response ("satisfied"). The other somewhat socially desirable variable, "It would be better for a pre-school child's emotional development if the mother didn't work outside the home," does not show a significant mean difference between the samples. Here the average telephone responses (5.12) are even slightly more in the socially desirable direction ("agree") than are the face-to-face responses (4.99).

These mixed results might mean that there is a mode effect in the predicted direction for questions which really do have a somewhat socially desirable response. Thus, with respect to the "satisfaction" question, people, simply because they are in the physical presence of another person, would rely more on the evaluative cues given by the other. Hence, they would be more likely to give a response that conforms with what they conceive the proper response to be, as reflected either in terms of their own perception of societal norms, or in terms of what they thought would be polite or pleasing to the other person. People may even feel that to answer that they are not satisfied with life in the presence of an interviewer who has gone to some effort to be pleasant would be an offensive kind of response. Thus, conditions specific to a face-to-face situation may well influence the answer given to the question, "how satisfied with life are you?" To the extent that they do, a mode effect is present.

However, there was not a significant mode difference on response means for the statement, "It would be better for a pre-school child's emotional development if the mother didn't work outside the home". In fact, contrary to my predictions, the telephone mean was slightly more in the socially desirable direction. One interpretation could be that agreeing with the statement was socially desirable for some groups, but not for others, an opinion held by half the experts on the informal survey. Another possibility, however, is that there could be

an interviewer effect operating in this case. Most of the interviewers for the survey were female and, therefore, could be working mothers themselves. Showing disapproval of mothers working by agreeing with the item (giving a higher value) may not have seemed tactful to people in the personal interview situation. Thus, findings of more or about the same level of social desirability in telephone responses as compared with face-to-face levels need not necessarily be seen as disconfirming the hypotheses. A mode effect could still be operating. Given the nature of the question and the circumstances of the interview (being questioned by a female interviewer), the socially desirable response may well be to disagreee. Some of these possible explanations will be examined in a later section.[16]

For the somewhat socially undesirable category, again, there are mixed results. The only variable with a significant difference in the predicted direction was the one asking whether the respondent agreed there was a chance that he/she or someone in the household would be unemployed next year. A slightly higher percentage of face-to-face respondents gave the more desirable response, that they or other members of their households had not been unemployed in the last three years (75.2% as opposed to 70.7% for the telephone), but the difference was not significant ($p=.38$).

On the other hand, the telephone respondents gave significantly more socially desirable answers on two items, whether they or other members of their households had received financial help from relatives in the past three years, and the number of hours of television watched in the past week. Thus, the findings on only one of the four somewhat socially undesirable items support the hypothesis. However, the items which do not support the hypotheses are all recall items. More will be said about this finding later. As for somewhat socially desirable items, only one of the two provides support.

[16]With the benefit of hindsight, it would have been beneficial also to examine other somewhat socially desirable variables which were not so potentially controversial. However, the choice of variables was limited.

Social Desirability Not a Factor

Two of the variables which are devoid of social desirability, according to the informal survey ("I don't think the Federal government cares much about what people like me think"; and "Awareness of ACCESS TV"), do not show significant mode differences, as predicted. However, for the question, "About how many hours in the past week have you personally watched ACCESS Network Television on cable channel 9?", face-to-face respondents reported watching significantly more hours of ACCESS TV than did telephone respondents. The hypotheses predicted no significant mode effect on means or proportions for questions without socially desirable answers unless recall was required. Hence, the findings on all three of the variables are supportive of the hypotheses.

In the case of the personal interviews showing a greater number of hours of watching ACCESS TV, the mode effect operating may have little to do with any social desirability inherent in the idea of watching ACCESS. As discussed earlier, people may simply put more of an effort into trying to remember accurately in a personal interview than they do in telephone interview because of the "demand characteristics" (Orne, 1962) of personal interviews. They would, thus, be expected to remember details better in personal interviews when the question deals with memory content. There may actually be a memory-based mode effect independent of the social desirability content of the particular item.

While the hypotheses for items neutral on social desirability were supported, it is still tempting to speculate about other interpretations hinted at by these results. Despite the supposed absence of any social desirability component in these variables, as suggested by the informal survey, it may be that there is still some kind of related effect on responses given. Thus, there is a slightly greater tendency for telephone respondents to agree that the Federal government doesn't care (mean of 5.17 as opposed to 4.94 for the face-to-face group). If there is confusion on the part of respondents as to who is sponsoring the survey, a not unusual situation, there may even be misunderstanding to the extent that people feel the EAS is a government survey of some sort. In such a case, it would be more threatening in a face-to-face situation to agree that the Federal government doesn't care than it would be over

the telephone. That is, the content of the question might interact with the characteristics of the sponsoring organization. Thus, social desirability can still be a factor even when the question content appears to argue against it.

Similarly, despite the opinion of those in the informal survey that social desirability is not a factor in saying that one is aware of ACCESS TV, demonstrating awareness as opposed to ignorance of any subject is certainly a desirable thing to do. It may be illustrated by the slightly higher proportion of those in the face-to-face condition claiming such awareness (76.7 versus 71.6). Or, it may be that watching and awareness of ACCESS TV are socially desirable activities, contrary to the opinions of the informal survey. Had people been asked their opinion of the social desirability of watching ACCESS as opposed to simply watching TV, and had these questions appeared adjacent to each other on the informal survey questionnaire, watching ACCESS may well have received a higher rating on social desirability.

Overall, then, first impressions of differences in desirable responses between telephone and face-to-face groups are, with qualifications, in the predicted direction. That is, the face-to-face group tended to give somewhat more desirable answers, with significant differences where the desirability or undesirability was not very clear. Exceptions to this pattern seem to occur for questions where the respondent was required to remember, where a particular reponse may have been desirable for some but not for others, or where the question format was complex. Thus, telephone responses on complex questions may be subject to misunderstandings which could be more easily dealt with in personal interviews.

B. Variability

Predictions were made in Chapter 2 for a somewhat greater amount of variability in responses for the telephone sample than for the face-to-face sample. This difference in variability would reach statistical significance on the items with "somewhat" socially desirable or undesirable responses. It was reasoned that simply reading response categories over the phone would emphasize endpoints on scales and, thus, increase the salience of these as appropriate response choices. Telephone respondents might also be less inclined to "converge"

their answers to what they felt was desirable to the interviewer, based on visual assessments. There was also the possibility that "prank playing" (providing extreme answers) is more likely in telephone situations where disapproval for such behaviour would not be as apparent.

The results on variances in Table 3 do not generally support these predictions. Only nine of the fourteen variables (those allowing for a choice of more than two responses) could be examined in this manner, and only four of these demonstrated the hypothesized outcomes. Significantly more variability was predicted for telephone respondents on the somewhat socially desirable or undesirable items, and for face-to-face respondents on recall items of neutral desirability. These predictions were only upheld on the neutral items, and on one somewhat desirable item, "How satisfied with life are you these days?". The other somewhat socially desirable or undesirable items, "It is better for a pre-school child's emotional development if the mother doesn't work outside the home" and "Number of hours of TV watched in the past week" showed significantly more variability for the face-to-face respondents. While the very socially desirable and very undesirable items were not expected to show significantly more variability for the telephone sample, two of the three items examined did, ("There should be stricter penalties for parents who do not properly secure their children in car safety seats" and "Punishments for impaired drivers are becoming too severe").

Again, there may be some value in trying to account for some of the discrepant and non-findings. If "hours watched TV" had been tested as a neutral variable, in terms of social desirability, a finding that face-to-face respondents showed more variability may simply reflect more of an effort put into remembering. The effect of having another person present has been observed to increase quantity and variety of response for open-ended questions, and a similar principle may be at work on these hours of television watched questions.

The greater variability of face-to-face responses on the question about mothers of pre-school children not working is not so easily explained, nor is the lack of a significant difference on the question, "There is a good chance that someone in the household will be unemployed next year." On the pre-school child question, one can, again, only speculate about possible interviewer effects deriving from using predominantly female interviewers or

Table 3

Variability in Responses of Face-to-Face and Telephone Samples

Variable[1]	F Statistic	p value	Telephone	Face-to-Face
#1-Very Socially Desirable				
Car seats for children	1.33*	0.02	1.70[2]	1.47
			2.89[3]	2.17
			.30[4]	.26
#2-Somewhat Socially Desirable				
Better for pre-school child if mother doesn't work	1.33**	0.033	1.7	1.95
			2.86	3.79
			.33	.39
How satisfied with life are you?	1.29*	0.07	1.39	1.23
			1.93	1.5
			.26	.22
#3-Neutral				
Federal government doesn't care	1.00	.51	1.73	1.73
			3.0	3.0
			.33	.35
Hours watched ACCESS TV	3.35**	.000	1.8	3.3
			3.4	11.2
			2.04	2.44
#4-Somewhat Socially Undesirable				
Hours watched TV past week	1.66**	0.0005	11.95	15.4
			142.8	237.2
			.92	1.02
Chance someone in household unemployed next year	1.12	0.234	2.41	2.29
			5.8	5.2
			.72	.82
#5-Very Socially Undesirable				
Impaired driver punishments too severe	1.73*	0.000	1.89	1.44
			3.6	2.1
			.83	.77
Okay to drive after 2-3 drinks	1.24	0.065	1.93	1.73
			3.7	3.0
			.84	.77

[1] Variability on dichotomous variables is not included here since a measure of variance provides no further information than is available in the proportion itself.

[2] Standard deviation.

[3] Variance.

[4] Coefficient of relative variation.

* Denotes significance at .05 level.

** Denotes significance at .05 level, but greater variance in face-to-face sample.

the possibility that the question is really only socially desirable for some people. It is possible that the unemployment item did not obtain expected results in variability because the issue was simply too sensitive, that is, unemployment may have been too likely an event for many people in Edmonton. In other words, perhaps it is an item belonging in the very socially undesirable category.

In the same way, the significantly greater telephone response variability on the car seats for children and the impaired driver punishment questions cannot be easily explained. It is possible that there was some difficulty understanding the impaired driver punishment question over the phone,[17] so this factor, in itself, could produce more variability in response. However, there is nothing to indicate that the "car seats for children" item was similarly difficult to understand.

In any case, the hypothesis that there would be more variability for telephone respondents on somewhat socially desirable or undesirable items has received limited support. There are probably other factors at work here which have not been taken into account. Several have been briefly discussed, but others might include such influences as memory requirements of the question, question complexity, interaction of item characteristics with sex of the interviewer, and misclassification of items with respect to degree of social desirability.

C. Nonresponse

It was predicted that nonresponse ("don't knows" and "missing" combined) would be greater in all cases for telephone respondents than for face-to-face respondents. The primary reason for this pattern would be that people feel more "compelled" to give an answer in personal interviews because of the salience of the physically present Other and because of the accompanying norms of reciprocity, politeness, and the attendant nonverbal attentiveness cues. In a way, then, simply responding may be seen as a socially desirable behaviour which would be expected more frequently in face-to-face situations. A related rationale for expecting less nonresponse for face-to-face interviews is the greater effort people would be

[17]Evidence that comprehension difficulties occurred for this item is presented in the next chapter.

likely to make in remembering information for someone who is actually present. Thus, there would be fewer "don't knows" for answers, as well as more details given by respondents for open-ended questions.

Although the pattern of non-response reveals that it is generally greater for telephone respondents, the difference is only significant on four variables (Table 4). For the thirteen variables considered, all but two show more nonresponse in the predicted direction, with neither of the exceptions showing significant differences. The nonsignificant exceptions are "car seats for children" and "okay to drive after 2-3 drinks."

The variables which do show significant differences in the predicted direction (i.e., more nonresponse for the telephone sample) were, "contributed to Food Bank last three years", "received financial help from relatives", "been unemployed past three years", and "received welfare past three years." More will be said in the next chapter about the fact that only these behavioural items showed significant differences by mode.

Overall, then, the hypotheses concerning mode effects on nonresponse received only limited support. On the other hand, it could be said that there were no contradictory findings, that is, no significant differences in the direction opposite to that predicted.

Table 4

Proportion of Nonresponse for Face-to-Face and Telephone Respondents

Variable	Telephone (N=116)	Face-toFace (N=464)	p Value (from chi square)
Rating #1 - Very Socially Desirable			
1. Car seats for children	0	1.1	.57
2. Contributed to Food Bank last 3 years	2.6	0	.006*
Rating #2 - Somewhat Socially Desirable			
1. Better for pre-school child if mother doesn'twork	2.6	0.4	.09
2. How satisfied with life are you?	0.9	0.2	.86
Rating #3 - Neutral			
1. Federal government doesn't care	4.3	1.7	.18
2. Hours watched ACCESS TV	.9	.4	1.00
3. Aware of ACCESS TV**			—
Ratins #4 - Somewhat Socially Undesirable			
1. Received financial help from relatives past 3 years	1.7	0	.05*
2. Hours watched TV past week	0	0	—
3. Chance household unemployed next year	4.6	2.2	.33
4. Been unemployed past 3 years?	1.7	0	.05*
Rating #5 - Very Socially Undesirable			
1. Impaired driver punishments too severe	1.7	0.9	.76
2. Okay to drive after 2-3 drinks	0	0.8	.71
3. Received welfare past 3 years	1.7	0	.05*

* Denotes significance at .05 level.

** Because of the way the questionnaire item was phrased, there would be no cases of nonresponse for this variable.

V. Further Analysis

It was decided to look further at some of the anomalous findings to see whether explanations or clarifications might emerge from possible interacting or suppressing effects from other variables. In other words, if a mode effect had been expected but did not seem to be present, what might account for the lack of such effects? Or, if it seemed that the interviewing mode was having an effect where one was not expected, what other factors might be responsible?

There is not much agreement in the literature on factors which might interact with interviewing mode. Phillips and Clancy (1972) found that item desirability response effects were greater for women than for men. But others have not detected interaction effects involving sex, age, education, or income (Gove and Geerken, 1977, cited in de Maio, 1984). These studies, though, were not examining comparisons by mode of interviewing. One study which did attempt to see whether mode differences in response patterns were greater for some groups than others was that of Aneshensel *et al.* (1982) which found that discrepancies between levels of reporting of depression symptoms face-to-face and by telephone were greatest for Blacks and Hispanics. Groves and Kahn (1979) also found method effects to be greater for nonwhites than whites, but they were not looking at social desirability specifically. Furthermore, they had expected interactions between both education and age and interview mode, but did not find any.

The potential interactive effects of sex, age, and education are examined below, since these variables are often important sources of interaction effects in other subject areas. Besides, there are grounds for thinking that age could interact with mode because of different experiences older people may have had with telephones. Also, it is reasonable to see whether results on interaction effects have been inconsistent because the degree of social desirability has not been examined in previous tests.

A. Item Desirability Hypotheses

Socially Desirable

We had predicted but did not find a significant mode difference in agreement or disagreement with the statement, "It would be better for a pre-school child's emotional development if the mother didn't work outside the home." Since the respondents in the informal survey had rated agreement with the statement as somewhat socially desirable, it was expected that the personal interview group would agree significantly more. They did not, perhaps because this response was really not socially desirable for all types of people, a possibility noted by half of the experts.

Bivariate regression of this opinion variable on mode of interview showed no significant effect for mode ($b = .1347$, $p = .50$), with face-to-face coded as 1 and telephone coded as 0. The variable was then regressed on both mode and sex of the respondent (scored 0 for men, 1 for women) to see whether controlling for sex might uncover the effect of mode. Sex was found to have a significant net effect ($b = -.5125$, $p = .0012$, two-tailed test). Men were more likely to agree with the statement than women. More agreement by men is to be expected, given current gender role attitudes. But interviewing mode still had no effect controlling for sex of the respondent ($p = .4920$).

An interaction term for sex and mode entered into the regression equation was not statistically significant. Thus, further analysis has not uncovered support for the original hypothesis. It may be that agreement is very, rather than somewhat, socially desirable for men. A mode effect for men would not, therefore, be predicted because nonverbal evaluative feedback cues are unnecessary. But the informal survey indicates that most men (seven out of nine, including interviewers but excluding experts)[11] felt agreement was only somewhat socially desirable.

---

[11]For the present purposes, expert opinion is ignored as it was supposed to capture opinions of what others thought, not what the experts themselves felt personally about the items.

The effects of education and age on the variable were also examined in a regression equation (including mode and sex). There was a main effect for years of education ($b = -.1203$, $p = .0000$), and age ($b = .029$, $p = .0000$), but no interaction effect between either education and mode or age and mode.

Thus, interpretation of these unexpected findings on the variable has not been clarified by further data analysis. The subject matter of the questionnaire item may simply not have any clear social desirability component because of changing mores, individual differences, or confounding effects of interviewers' sex. In any case, the results do not support the original hypothesis of significantly more socially desirable response in personal interviews for "somewhat" socially desirable items. However, the findings for the other somewhat socially desirable item ("how satisfied with life are you?") do support the hypothesis. Future tests of this theory should include a larger number of somewhat socially desirable items.

Neutral Desirability

For variables considered to have no social desirability component in a particular response, no significant mean or proportion differences by mode were predicted, unless recall was required. Therefore, when the face-to-face sample reported watching significantly more ACCESS TV, this was considered an anticipated finding. Nevertheless, a further look at one possible interaction effect would be interesting.

Perhaps, as noted earlier, there might actually be a social desirability component to the reporting of watching ACCESS TV. If so, our theory would still predict more hours reported in face-to-face interviews. The t-test did show a probability of .021 (one-tailed test) of finding such a difference between the means. Controlling on sex, age and education in a regression analysis did not change the relationship. A two-way analysis of variance was then done, comparing means on hours watched by mode for four different levels of education. Neither the main effects for level of education or mode were significant, nor were the

two-way interactions.[19] But the patterning of the means for the groups is illuminating (Table 5).

For the three lower education levels, there are consistently higher means reported in the face-to-face situation, while for the group with the highest education level, there is practically no difference between modes. This pattern may be evidence of a difference by education level in how socially desirable it is to report watching ACCESS. People with a high level of education do not report watching more educational television when interviewed face-to-face, possibly because they take watching educational television for granted. People with lower levels of education may, however, be attempting to present a more favourable impression of themselves in personal interviews.

We may have here an interesting case of an idea proposed by De Maio (1984) to link the two primary conceptions of social desirability . It may be the case "that certain items are endorsed (artifactually) just because they are desirable, *especially by those who need social approval*" [italics mine] (p. 271). For those with less education, responses to this item may reflect both personality and item-based social desirability. It is possible,then, that saying one watches a lot of ACCESS TV is, in fact, a somewhat socially desirable response, at least for some types of people. The experts and others who, on the informal survey, rated the item as not having a social desirability component were themselves perhaps showing biases produced by their own high level of education. A better future test of mode effects would require more items more truly neutral in desirability for all respondents.

Even if the ACCESS TV item was misclassified in that saying one watches a great deal of ACCESS is actually somewhat socially desirable, results could still be considered to support the theory. The people for whom it is probably somewhat socially desirable to say they watch ACCESS, those with less education, are the ones most affected by the face-to-face mode of interviewing.

---

[19]Absence of a statistically significant interaction effect for level of education and mode could be due to the relatively large variance around each mean (see Table 5 for standard deviations).

Table 5

Amount of ACCESS TV Watched Past Week by Mode and Level of Education

Level of Education	Means (Coded in Minutes), N of Respondents, and Standard Deviations	
	Face-to-Face	Telephone
1 (0 through junior high)	84.00 (35) 142.7	39.23 (13) 72.9
2 (High School and incomplete H.S.)	94.04 (171) 270.3	61.84 (49) 133.7
3 (Technical and incomplete university)	75.00 (145) 140.7	41.38 (29) 61.8
4 (University and beyond)	69.88 (82) 129.5	67.50 (22) 123.2

Socially Undesirable

For the somewhat socially undesirable category a variable about watching television (of any kind) showed significantly more hours watched in the face-to-face sample, contrary to what was predicted. A t-test had shown a probability of .048 (one-tailed test) of obtaining such a difference between the sample means. Regressing hours of television watched on mode, sex, education, and age showed significant effects for education ($b = -43.1802$, $p = .0001$) and age ($b = 5.5555$, $p = .018$), but not for mode ($p = .1434$).[20] So, older people reported watching more television, and the more educated watched less. There were no interaction effects between mode and sex, age, or education when interaction terms were added to the regression equation.

Another somewhat socially undesirable rated variable, "Have you or anyone else in your household received financial help from relatives in the past three years?" showed significantly more socially desirable answers ("no") in the telephone sample when the opposite result had been predicted. To see if the frequency of a "no" response differed in modes by sex, crosstabs showing the relationship between mode and response were set up controlling for sex. Sex was not found to interact with mode. Similar results were obtained controlling for age and then years of education and, although increasing age and years of education reduced mode's effect on the proportion of "no" responses, the effects were not significant.

A logistic regression was tried in which the probability of saying "no" to "received financial help from relatives" was regressed on mode, age of respondent, years of education, and sex in a logit model. The following coefficients/standard errors[21] were obtained for the independent variables:

---

[20]The bivariate regression of hours of television watched (coded in minutes) on mode (telephone coded 1, face-to-face coded 0) had only shown a significance level of .15 ($b = -132.4$), but mode will be considered here to have a significant effect because of the t-test results. Discrepancies in significance levels are due to use of different types of variance estimates in the two procedures and employment of a two-tailed test in the regressions.

[21]According to Walsh (1987), these coeff./s.e.'s (the regression coefficient divided by its standard error) should be considered analogous to a t value, so that useful variables in the equation should have a coeff./s.e. value of 2.0 or more.

1. Mode = 2.67987
2. Age = 6.57223
3. Education = -1.77458
4. Sex = .04698

Mode and age had important effects on the probability of saying "no" with telephone respondents more likely to say no (telephone mode was coded 1). Interaction terms between mode and age and mode and education were both tried out in the logistic regression, but mode's effect effect on saying "no" did not differ by respondents' ages or levels of education.

Without the interaction terms in the equation, then, with the other independent variables controlled, mode still has a significant effect, and it is in the opposite direction from that predicted. People provide the undesirable "yes" response more often face-to-face. One explanation which springs to mind is that people are being asked to remember something (over a three-year period) so we could expect memory to be better face-to-face. They are also being asked whether they or any other members of their households received financial help from relatives. They are not only being asked to remember something concerning themselves, but concerning other people, too. The easiest and least inconvenient response to make is "no" and this "lazy" response tendency may override the social undesirability factor when respondents are asked about past financial help over the telephone.

The theory may be too simple in that it proposes the predominance of social desirability as a factor in people's answers. Motivation to respond with thorough information may simply be a more overriding concern face-to-face, while telephone respondents "take the easy way out" on questions requiring mental effort. The fact that telephone responses appear to be more socially desirable may simply be coincidental, then, in such a scenario.

It could be argued, though, that the same factors should be operating in responses people gave to the "received welfare in the past three years" question where face-to-face respondents gave the socially desirable "no" response slightly more often. However, receiving welfare has been rated as very socially undesirable, so, in this case, we did not expect a mode difference. Furthermore, it is likely that whether someone in the household has received

welfare will be more easily remembered than receiving financial help because welfare is more traumatic. A final difference is that whether anyone had received welfare is an event which is subject to verification, whereas receiving financial help from relatives is not. Immediate interactional considerations of favourable presentation of self may be outweighed by the chance of being caught in a lie.

The other somewhat socially undesirable item with unexpected results asked whether the respondent had been unemployed for up to a month in the past three years. But the difference between samples was not significant, and respondents actually gave the undesirable response "yes" more often in face-to-face interviews.

A crosstabulation controlling for sex was done which indicated that women were giving the socially desirable "no" response more often face-to-face while men did so more often on the telephone. Crosstabulations were also tried controlling for level of education and age, but these did not appear to indicate any interaction between mode and either of these independent variables. To verify the relationship between mode and sex, a t-test was done comparing women's responses by mode. This showed the the difference in proportion of "no" responses for women to be statistically significant by mode ($t=1.75$, $p=.04$, one-tailed test).

It seems surprising that for this question about being unemployed, women would be giving the socially desirable "no" response more often in face-to-face interviews than men. One expects that it would be more socially undesirable for a man to be unemployed than it is for a woman, so men should be less willing to admit to having been unemployed in front of another person.

A logistic regression was then tried with the probability of saying "no" to "Have you personally been unemployed?" (in the last three years) regressed on mode, age, years of education and sex in a logit model. When the coefficients/standard errors for the independent variables were obtained, the results showed that age, education, and sex were all important predictors of the probability of saying "no." The coeff./s.e. for each of the independent variables in the equation were:

1. Mode = -.38652

2. Age = 6.04836

3. Education = 2.60206

4. Sex = 2.03812

Thus, women (coded 1) were still more likely to say "no" than men were, reaffirming the findings from the crosstabulations even when the effects of mode, age, and education were controlled in the same equation. Such a surprising finding might be explained by the possibility that women answer "no" more because they tend not to perceive themselves as "unemployed".

Because a greater proportion of the women in the face-to-face sample "kept house" than in the telephone sample (17% versus 10%), an interaction term between mode and sex was added to the logistic regression equation along with the other predictor variables, mode, age, education, and sex. A coefficient/standard error of -1.92000 for the interaction term resulted, a figure large enough to indicate that the interaction between mode and sex was important in predicting the probability of saying "no". Since this result still did not make sense intuitively, I looked at whether the findings could be explained by effects of respondent's employment status, that is, whether someone was employed or was keeping house.[22] When this variable was controlled for, the only important original predictor of a "no" response which remained was age. Both sex and the interaction between sex and mode were no longer influential (the coeff./s.e.'s for each dropped to 1.11605 and -1.10149 respectively).

What the above shows is that, ultimately, the only variables which seemed to be useful at all for predicting the probability of a "no" answer to whether the respondent had been unemployed in the past three years were the respondent's age and employment status. Mode of interviewing was not important, nor was sex, when employment status was taken into account. Whether the lack of any mode effect when such effects had been predicted means that the content of the item is not somewhat undesirable as expected is not clear. There

[22]This was tested by adding a dummy variable for the various employment statuses to the same logistic regression equation. Each category was found to have an important effect on the probability of saying "no."

were slightly more "no" responses face-to-face, but probably the combination of having a slightly higher proportion of homemakers in the personal interview sample as well as the likelihood that unemployment is not so undesirable for women has kept the difference between samples from attaining significance.

The fact that an earlier regression showed that being female had an important positive effect on the probability of answering "no" is an indication that unemployment could be socially undesirable for some, but not others, at least when considered in relation to one's self in answering a question. This item, then, may not have been a good choice for testing the hypothesis.

Very Socially Undesirable

In the very socially undesirable category, there was a significant difference between telephone and face-to-face means on the "Punishments for impaired drivers are becoming too severe" question. Because of the strength of norms about the impropriety of drinking and driving in our society, it had been predicted that there would not be a significant difference between samples. The significantly higher (more socially undesirable) telephone sample mean was surprising, because the samples had not differed on a related question, "It's okay to drive soon after drinking two or three alcoholic drinks."

Bivariate regressions of the impaired driver punishments question on mode, on sex, on education, and on age, showed the first three of these variables, mode, sex, and education, to have significant slopes. People in the telephone sample agreed more, men agreed more than women, and those with more education disagreed more with the statement. There were no significant interaction effects involving either age or sex, and mode. However, introducing an interaction term between education and mode into the equation (along with the other variables) yielded a significant slope for the education/mode interaction ($b = -.112$, $p = .018$). This meant that the effect of the interviewing mode varied with the respondent's level of education.

To examine this effect more closely, respondents were divided into two groups (up to completed high school, and education beyond high school). The means of these two groups were compared, by mode, using an analysis of variance (Table 6). The means for the high education respondents were similar in both interviewing modes, and were little different from the mean for the low education respondents interviewed face-to-face. The anomaly among the four groups was the much higher (least socially desirable) mean for the low education group interviewed by telephone.

Telephone respondents, whatever their level of education, do not hold less socially desirable opinions about drinking and driving (Table 7). But level of education did make a difference in their understanding of the "punishment" question, if it was asked over the telephone. The form of this question is more complex, in that it asks for level of agreement with a statement endorsing a socially undesirable stance. The question on driving after drinking does this, too, but is simpler to understand. The simple key phrase, "It's okay to" highlights what the respondent is being asked to agree with. The impaired driver question, on the other hand, uses less familiar vocabulary, (e.g., "impaired") and has a double negative structure, if one disagrees. The respondent must think about punishing someone, the meaning of "impaired driver," and whether punishments for such people are severe or not, and then try to remember exactly with what it was he or she was agreeing or disagreeing. The combination of less education (and, hence, less experience with complicated verbal statements), plus the absence of the interviewer to provide clarification through questions, gestures or other communication methods, can lead to a response which, on the surface, appears socially undesirable.[23]

---

[23]An independent test of the difficulty of understanding this particular question was provided by the respondents in the informal survey. Of the twenty-six non-experts, eleven rated agreement with the impaired driver punishment question, which is rated as a socially undesirable response, as very, somewhat, or neutral in social desirability. This suggests some misunderstanding. Agreement with the "okay to drive after two or three drinks" question, on the other hand, was mistakenly rated by only six of the non-experts. Experts did not exhibit confusion on either item.

Table 6

Means (by Mode) of High and Low Education Groups

on Impaired Driver Punishments Question

	Type of Sample	
	Face-to Face	Telephone
Education		
Low	1.86 (221)	2.75 (61)
High	1.78 (239)	1.71 (51)

Table 7

Means (by Mode) of High and Low Education Groups on Driving After Drinking Question

	Type of Sample	
	Face-to-Face	Telephone
Education		
Low	2.31 (221)	2.25 (63)
High	2.20 (239)	2.27 (51)

B. Variability

There was no further statistical analysis done for findings contrary to predictions regarding either variability or nonresponse. For the latter, it is difficult to carry out further analyses because the proportions of nonresponse are so very low. And, some ideas relevant to nonresponse and variability have already been discussed above.

It had been hypothesized that variability would be significantly greater for telephone respondents on those questions which had somewhat socially desirable or undesirable answers. Face-to-face responses would show significantly greater variability on items requiring recall, but only if there was no social desirability component. However, it appears likely that, in some instances, considerations other than social desirability come into play to influence the type of answer one gives. For example, being asked to recall details may override considerations of social desirability. This requirement may explain why there is significantly more variability in response for the face-to-face sample on the two variables measuring television watching. Face-to-face respondents may be putting more of an effort into remembering and, thus, come up with a greater variety of responses. Personal interviews seem to increase quantity and variety of response for open-ended questions (Groves and Kahn, 1979) and a similar principle may be at work on these hours of television watched questions.

Other contrary findings on variability are not so readily explained, however. The "impaired driver punishments" question may have exhibited more variability on the telephone because of its complexity, and there is a slight possibility that the car seats for children item posed similar difficulties. In the case of the greater variability in face-to-face responses on the question about mothers of pre-school children not working, it may be that an interviewer effect is operating. Respondents who are faced with female interviewers may feel conflicting pressures as to how to choose the "right" response. Furthermore, the fact that there may not be uniform opinion on the social desirability of agreeing with the statement may contribute to variability as well.

Possibly, variability is a measure which is not as greatly influenced by considerations of social desirability as assumed. Or, factors such as requiring recall or asking a complex

question have played a greater role than expected. Some of the same difficulties encountered earlier in assessing differences between samples on means or proportions, such as the possibility that an answer was socially desirable for some, but not others, or that items may have been misclassified, could also lead to unexpected findings on variability. In any case, the hypothesis that there would only be more variability for telephone respondents on somewhat desirable and undesirable variables has received very limited support. There appear to be other factors at work here which were not taken into account at the prediction stage.

C. Nonresponse

The results from the examination of nonresponse were not subjected to any further statistical analysis as there were so few cases and findings were generally in line with the predicted direction of more nonresponse for the telephone sample. However, differences in nonresponse were not significantly greater for telephone respondents as predicted except for four variables. The two items for which there was slightly more nonresponse in the face-to-face group ("Car seats for children" and "Okay to drive after 2-3 drinks") showed a difference in direction on nonresponse which could have occurred by chance alone.

The four items which did show significantly more nonresponse in the predicted direction (i.e., for telephone respondents) were all inquiries about a behaviour rather than an attitude or an opinion, with the face-to-face group showing 0% nonresponse on these behavioural items. All four significantly different questions came from the same section of the questionnaire, so it may be important to know what effect the grouping of these items together may have had. Furthermore, the introduction to this section, finding out about "how changes in the Alberta economy have affected your household", may have influenced the way people responded.

Significantly fewer cases of "don't knows" and nonresponse for personal interview respondents do not necessarily indicate that the face-to-face group were more concerned with presenting a socially desirable image of themselves, either, since telephone respondents gave the socially desirable answer, "no" more frequently to the item asking if they or anyone in

their household had received financial help from relatives. Here, then, for these four items, it would be valuable to test whether face-to-face nonresponse is less because behaviours are the subject matter (with the implication to the respondent being that responses could be checked for accuracy) or because the subject matter was the recession, a topic of more immediate concern to people in Edmonton than other subjects might be at the time. It may even be that more worrisome items create different response patterns in interaction situations, so that the presence of another person when such items are discussed elicits more response.

The question asking the chance of being unemployed next year, on the other hand, did not produce a significantly different pattern of nonresponse. Yet, the subject matter related to the recession as it did for the behavioural items which did exhibit patterns of nonresponse as predicted. There is also the possibility that, even though the subject matter of this chance of unemployment item is very similar to the four significantly different variables, the fact that an opinion is asked for rather than a behaviour or factual item, may be responsible for the different pattern in nonresponse. It should be noted, though, that the usual pattern of somewhat greater nonresponse for the telephone sample is maintained.

Looking for explanations of unexpected findings regarding nonresponse based on an item's desirability or undesirability does not reveal any clear pattern. There is an exception to the general pattern of more nonresponse on the telephone in both the very desirable (car seats for children) and in the very undesirable category (okay to drive after 2-3 drinks), although neither is significant, so social desirability of the item itself is probably not a factor in nonresponse for these two items. The fact that three of the four significant differences in nonresponse occur in the undesirable items does not necessarily mean very much either since a) there happen to be more undesirable items examined and, thus, more chance of finding differences in this area and b) all of the significant differences in nonresponse are on variables concerned with recession-related behaviours.

Thus, to summarize, there is a general tendency for more nonresponse among telephone respondents, but differences in nonresponse on our chosen variables are only significant when the question concerns behaviours or, specifically, hardships related to the

recession. To say anything further about the effect here of the question's subject matter would require examination of more items from the Edmonton Area Survey.

VI. Conclusion

A. Summary of Mode Effects

To attempt an overall assessment of whether mode of interviewing did have the predicted effects on responses to questions asked on the 1988 Edmonton Area Survey, it would be useful first to review these predictions. It was proposed that the nature of communication in interviews would differ between face-to-face and telephone situations in ways that would relate both to a person's choice and amount of response on certain kinds of questions, as well as to his or her decision whether to answer questions at all. Because the physical presence of another human being seems to produce a kind of solidarity or accommodation between the individuals, it was proposed that how one was interviewed would particularly affect people's responses on questions which had answers which were socially desirable or undesirable, since how a person answers such questions affects the other's evaluation. The face-to-face situation makes any such evaluation more immediate and important to the respondent.

However, for questions subject to the influence of social desirability, it should only be on those topics which have somewhat socially desirable or undesirable answers where interviewing mode will cause response differences. While nonverbal evaluative feedback cues help respondents choose answers to such questions, some topics arouse such strong feelings on the part of an individual that asking whether he or she does a certain thing or agrees with a certain statement makes the evaluative response of another person irrelevant. That is, the individual has a strong ego involvement with answering the question in a particular way due to a strong societal norm about the propriety of a certain response, or to a strong personal sense of involvement or commitment to answering in a specific way due to the individual's personal experiences. For such questions, the mode of interviewing should not affect the answer one gives. On those items for which there was neither a socially desirable nor undesirable response, there should be no difference by mode in responses unless memory recall was required. If recall was required, face-to-face respondents would be motivated to produce

more thorough, and thoughtful responses. These predictions were to be tested through examining differences in means or proportions on responses between samples differing by mode of interviewing.

Some predictions were also made regarding the relative variability of responses in the two interviewing modes. Significant differences in variability would occur where certain responses to items were only somewhat socially desirable or undesirable; because the evaluative presence of the Other in face-to-face interviews would have that much more of a channelling effect on responses. Without this evaluative Other, the telephone answers would be much more varied since the respondents would be missing the information which nonverbal cues provide to guide choice of the desirable response option.

It was also predicted that for items requiring recall which did not have a socially desirable or undesirable response, there would be more variability in face-to-face answers. There would still be a social motivating force in personal interviews based, at least in part, on nonverbal attentiveness cues which would promote more effort and thoroughness in responding. The same rationale applied to predicting significantly more nonresponse for telephone respondents on all items.

Testing for variability was essentially done by means of an F-test of variances. For proportions of nonresponse, predictions were simply for significantly less nonresponse in the face-to-face mode because of the social desirability of giving some answer in the compelling physical presence of the person asking the question. Testing involved comparison of proportions of nonresponse in both samples with significance of differences in proportions being shown by Fisher's Exact Test.

The predictions made for means and proportions, variability, and nonresponse due to mode of interviewing are summarized again for convenience in Table 8.

Success of Predictions

It would be useful also to summarize how the chosen variables fared on these tests, as a way to evaluate the importance of social desirability in producing mode effects (Table 9). If

Table 8

Predictions

Category	Mean or Proportion	Variability	Nonresponse
1. Very Socially Desirable	Same in both modes	Same in both modes	Significantly more for telephone
2. Somewhat Socially Desirable	Significantly in more socially desirable direction for face-to-face	Significantly more for telephone	Significantly more for telephone
3. Neutral Desirability	Same but significantly different if recall required	Same but significantly more for face-to-face if recall required	Significantly more for telephone
4. Somewhat Socially Undesirable	Significantly in more socially desirable direction for face-to-face	Significantly more for telephone	Significantly more for telephone
5. Very Socially Undesirable	Same in both modes	Same in both modes	Significantly more for telephone

Table 9

Summary of Findings

Variable	Mean or Proportion	Variability	Nonresponse
1. Very Socially Desirable			
Car Seats for children	a	c	b
Contributed to Food Bank	a	___*	a
2. Somewhat Socially Desirable			
Better for pre-school child if mother doesn't work	b	c	b
How satisfied with life are you	a	a	b
3. Neutral in Social Desirability			
Federal government doesn't care	a	a	b
Hours watched ACCESS TV	a	a	b
Aware of ACCESS TV	a	___*	___**
4. Somewhat Socially Undesirable			
Received financial help from relatives past three years	c	___*	a
Hours watched TV past week	c	c	b
Chance someone in household unemployed next year	a	b	b
Been unemployed last 3 years?	b	___*	a
5. Very Socially Undesirable			
Impaired driver punishments too severe	c	c	b
Okay to drive after 2-3 drinks	a	a	b
Received welfare past 3 years?	a	___*	a

Explanation of Codes:
a = predictions confirmed
b = non-significant findings when significant differences were predicted
c = significant differences when none were predicted

---

* Denotes a variable having a dichotomous response which cannot, therefore, be compared for variability between samples.

** Nonresponse was not measured here as people were simply asked to name ACCESS Network activities of which they were aware. Thus, awareness was coded "yes" or "no", so if the respondent did not mention that he or she was aware of ACCESS TV, the response was coded "No", not NR (nonresponse).

all predictions had been fulfilled, results on each variable would be coded as "a". Obviously, there were many unexpected findings, but Table 9 is not particularly revealing regarding the complexity of possible explanations for these non-predicted findings.

If we examine the findings in Column 1 (Means or Proportions), for example, we can see that predictions on five of the fourteen variables were different from expectations. Furthermore, predictions for the very socially desirable and very socially undesirable items were more successful than were those for the other categories, which in itself is not surprising since these were, in effect, predicting a fairly certain outcome (i.e., lack of a mode effect). The one "c" in categories 1 and 5 was the "Impaired driver punishments" question. Here, although the direction of difference between modes fit the general predictions, the difference was significant when a difference was not predicted. In Chapter 2 the possibility was mentioned that mode differences on neutral items might emerge when not expected because of such factors as complexity of a question. It was not anticipated, however, that the same effect might occur on a very socially undesirable item. Nor was it realized at the time the "punishment" question might be difficult to understand.

The other exceptions in this column are more problematic for the prediction that items with somewhat socially desirable or undesirable answers would be most affected by mode and that without a social desirability component, mode would have little effect. Of the nine variables in these three categories (ratings 2, 3, and 4), four showed unexpected mean/proportion results. However, agreement with the "Better for pre-school child if mother doesn't work" item is probably not universally accepted as being somewhat socially desirable, is probably controversial, and may be peculiarly susceptible to interviewer effects. Therefore, it may not be a good test for mode effects deriving from social desirability of the item.

The same criticism might be levelled at the "Have you been unemployed in the past three years?" item since men and women are likely to differ on both their conceptions of unemployment and on their views of the degree of undesirability of being unemployed. Another potentially confounding problem with the desirability assessment of saying "yes" to having been unemployed is that there is no way to reconstruct how the respondents in the

informal rating survey were evaluating a "yes" answer. They were asked how they would rate a "yes" response, but it would also be useful to know if they were picturing a man or a woman, an old person or a young person, themselves or people in general. With the benefit of hindsight, the mother working and unemployment items should probably have been replaced with less problematic items for a fairer test of the hypothesis.

The remaining exceptions, "Hours watched TV past week", and "Have you or any other members of your household received financial help from relatives?" are both questions asking the respondent to put a fair bit of effort into recalling details of his or her own behaviour as well as events occurring to other people. We might expect that responding to these would be done "better" in a face-to-face situation, even if there were no considerations of item social desirability operating, simply because of the respondent's desire to please the other. But it was predicted, in effect, that the social desirability of responding in a certain way would be more important than this accommodative aspect, i.e., that one's self image in terms of societal norms would be overriding. The exceptions provide some evidence that giving a thorough response (remembering better and, thus, pleasing the interviewer) may outweigh considerations of societal definitions of social desirability when a person is asked to recall behaviours or events, at least when the subject matter is somewhat socially undesirable.

Although the "Aware of ACCESS TV" and "Hours watched ACCESS TV in the past week" items showed results as predicted, I feel now that they may have done so for the wrong reasons. There probably is a social desirability component to showing awareness, regardless of what the informal survey results showed, and I feel that it is probably somewhat socially desirable to show such awareness. Thus, I should have predicted significantly more hours of ACCESS TV watched face-to-face because it is somewhat socially desirable to watch ACCESS, rather than because the item was neutral in desirability but required recall effort. Also, predictions should have been for significantly more face-to-face "yes" responses on the "Aware of ACCESS TV" question for the same reason. However, the latter question was asked in a way that would prevent demonstrating any mode effect. People were simply told that ACCESS Network is involved in various activities and they were asked to name ones of

which they were aware. Most people are only aware of the television service provided by ACCESS, not the other services, so they name this one. Thus, it is less a matter of giving a socially desirable response by citing ACCESS TV service than it is one of recalling what is both a neutral piece of information and just about the only piece of information on the subject of which anyone is aware. On the other hand, asking how many hours of ACCESS TV a person has watched does get at answers which have a social desirability component.

If the two variables concerning ACCESS TV are, in fact, behaving in this way, i.e., that the awareness variable is really a neutral recall question while the hours watched variable picks up a somewhat socially desirable dimension, then the results on the mean/proportion differences between samples may be more understandable. If the television service is the only familiar thing about ACCESS to most people, then we need not expect significant differences by mode. Similarly, we would expect significantly more hours of ACCESS watched to be reported face-to-face, as obtained, if watching ACCESS is somewhat socially desirable.

A final comment about the findings on mode effects for differences in means or proportions is that the examples demonstrating results in line with predictions are mostly attitude questions. The attitude exceptions, as noted, were those which probably had responses desirable to some but not others, or else were complex questions. Where an attitude does have a response considered by most people to be somewhat socially desirable or undesirable and where the item is easily understood, there does appear to be a mode effect on the response given. We can say either that the face-to-face mode seems to produce a more socially desirable response or, alternatively, that the telephone mode seems to produce a more socially undesirable response. With so few examples of non-problematic attitude questions having been tested, however, it can only be tentatively concluded that there is some support for the hypotheses.

The other items with unexpected results were all recall questions, "In the past three years have you or any other members of your household received financial help from relatives?", "Hours of TV watched in the past week", and "In the past three years have you personally been unemployed for up to a month?" Some might claim that they did not show

the mode effects expected because they concerned behaviours rather than attitudes. However, I feel that attributing differences to "behaviours" might obscure the effects of the dynamics of the interview situation in affecting response outcomes. On the other hand, seeing the response as "remembering" can help explain why people might even "remember" more "bad" behaviour in the face-to-face situation than they would on the telephone.

In a sense, then, for these questions, social desirability which affects the response given may be of a different sort. The determinant of the direction of the answer a person gives is not the societal-wide norm as it is for the very socially desirable or undesirable attitudes and behaviours. Rather, for the behaviour/recall items for which there is only some or no social desirability, the principle guiding the repondent's answer appears to be "please the interviewer" by providing a thorough answer.

I do feel, however, that these questions, though all ranked as "somewhat" socially undesirable, are not necessarily equally so. Thus, having been unemployed is probably "worse" than watching a lot of television or having had someone in the household receive financial help from relatives. There may be an indication that this is the case in the marginally greater proportion of "no" responses face-to-face for the unemployment question (which may well have been significantly greater if sex had not been affecting "no" responses so greatly).

Such generalizations may not be justifiable from so few examples, but they may be worth testing further another time. For such a purpose, it would be necessary to try to explain why behaviour/recall responses seem subject to different influences than attitude responses. There is always the possibility that what one has done is more real and undeniable to a person than is an opinion. At the very socially desirable or undesirable end of the continuum, acts may be well remembered because of the ego involvement at these extremes, and the interviewer will not have much influence on one's effort remembering these. In the [illegible] categories, attitudes are less ego-involving than acts, and therefore, more likely to be influenced by prevalent social norms in face-to-face reporting. The undeniability of an act, though, may force a person to feel obliged to report it when the Other inquires about it

directly.

Alternatively, one might argue that remembering itself is more important in the personal interview situation when speaking of acts or events in the respondent's life. This aspect of the social desirability inherent in the interview interaction may, coincidentally, make it look as if it is the very socially desirable or very socially undesirable nature of a particular response to a particular question, or the item characteristic, which is responsible for similarities in responses to attitude and behaviour questions when modes are compared.

In a way, support for these latter ideas can be gathered from Column 3 of Table 9 (Nonresponse). It had been predicted that there would be significantly more nonresponse on all variables for the telephone respondents. However, all the cases where "a" appears (prediction upheld, significant or nonsignificant as predicted, direction as predicted) are items on which the respondent had to recall information. He or she did so significantly "better" in face-to-face interviews across all categories of social desirability. This better recall is evidence that social desirability may be operating differently as a factor influencing responses on behaviour/recall items. In other words, the respondent's motivation appears to have a different basis on recall as opposed to attitude items.

For the b's in Column 3, all are attitude items except for the Hours of TV watched item. The lack of a significant mode difference on this behaviour item may not invalidate the above ideas about greater respondent motivation to recall face-to-face. The question format was a bit different in that respondents could answer "0" hours as an alternative to not answering, so this factor may be responsible for the lack of a significant mode difference in nonresponse. There still appears to be a basic division in nonresponse patterns between attitude and behaviour items, with only the behavioural items showing mode differences in nonresponse. For nonresponse, then, it appears that hypotheses were only supported for behaviour/recall items.

It is likely that at least some of the unexpected findings on differences or lack of differences in variability between modes relate to problems already discussed in the foregoing sections. For example, the "Car seats for children" item showed significantly more variation

in the telephone sample when no difference was predicted. The question itself may have been a bit complex as it asked respondents to agree or disagree with the statement, "There should be stricter penalties for parents who do not properly secure their children in car safety seats". It is similar to the double negative on the impaired driver punishment question and may have led to a significantly greater variety of response (indicative of confusion) on the telephone. The impaired driver question also showed the same pattern of unexpected variability which was significantly different.

For the remaining two unexpected variability findings, one is a behaviour recall item ("Hours watched TV past week") and rationales for why variability might be greater in face-to-face situations have already been suggested. The remaining item, "Chance someone in household unemployed next year" showed more variability on the telephone, as predicted, but the difference was not significant. I have no explanation for this finding other than the possibility that the item could have been very rather than somewhat socially undesirable. In this case, a difference in variability would not have been expected.

To summarize, variability was predicted to be an important indicator of mode effects in that it would be primarily dependent on the distinction between the "very" and "somewhat" levels of social desirability and undesirability. It seems, though, that there is no obvious basis for the patterning of results. Only nine of the fourteen variables could be examined, however, and because of possible misclassifications and unanticipated question complexity, hypotheses about mode and variability probably did not receive a fair test.

Some of the same rationales for unexpected results which spring to mind have already appeared in connection with discussion of mean/proportion as well as nonresponse unexpected findings. Complications in predicting variability have arisen where question format may have been complex (childrens' car seats, impaired driver punishments), where there was a possibility of a stronger interviewer effect because of identification of the interviewer's sex with question content (pre-school child item), where motivation to recall accurately appears to override undesirability (TV watching), and perhaps where the seriousness of the issue to the respondent overcomes concern with self-presentation (chance of someone in household

being unemployed). The present analysis has not been able to support the hypothesis that mode of interviewing itself will cause differences in variability when an item has a "somewhat" socially desirable or undesirable response. Either other factors besides desirability are more important in influencing variability or, as has already been suggested, there could be problems with some of the evaluations of desirability. Variability may simply not be very responsive to mode effects caused by degree of desirability.

The overall conclusion about social desirability-related mode effects which can be drawn from these data is that they are there, but that they do not operate in quite the ways that had been hypothesized. Societal norms about the social desirability of certain responses to particular questions do seem to make a difference in answers given, which differs by mode. However, the effect is limited to attitude questions where there is consensus that an answer is somewhat socially desirable or undesirable. Thus, hypotheses were supported, but only for attitude questions and only for those which were not controversial.

Mode appears to operate in influencing one's answer to behaviour/recall questions in a different way. Here, how one answers (in terms of the desirability of the response itself) is not so much related to the content desirability of the question, but to the desirability of affiliative behaviour in the interview interaction. One could say that mode still makes a difference in how one responds to questions because of social desirability, but it is a different kind of social desirability, more of a responsiveness to the Other evidenced by greater effort to provide detail in answering. Presumably, this effect is based on the rewarding nature of nonverbal attention showing feedback cues provided by the interviewer. This effect was surprisingly strong in comparison to that of item desirability.

Reevaluation of the Role of Nonverbal Cues

Given the results obtained in this study, it might be useful to reevaluate the ideas about how mode of interviewing and nonverbal communication might affect answers to questions having a social desirability component. The presence of nonverbal cues in the face-to-face situation is still considered to be the underlying explanatory factor in mode

differences on these items. However, the primary difference from the earlier formulation is that different types of nonverbal cues are predicted to "kick in" at different times. In other words, one type of cue is proposed to take over from another type as being more important in guiding responses, say at a certain level of social desirability and depending on whether the item is attitudinal or behavioural.

Because there were more obvious patterns to the mode differences obtained in the mean/proportion and nonresponse measures, as well as fewer usable variables for the variability measure, discussion concerning changes in understanding how nonverbal cues could affect responses will be based on findings in Columns 1 and 3 in Table 9. Thus, at the very socially desirable or very undesirable ends of the desirability continuum, it would appear that nonverbal evaluative cues regarding item desirability are not particularly relevant to the response a person makes. This irrelevance is shown by the absence of mode differences, regardless of whether the item is attitudinal or behavioural. The desirable response is given regardless. It seems likely that item characteristic desirability based on social norms is the relevant factor in the response given and that nonverbal evaluative cues, if present, would simply constitute "noise" in the system.

On the somewhat socially desirable and undesirable items, it looks as if the nonverbal cues responsible for mode differences might be different depending on whether an attitude or behaviour is the subject of the question. For attitudinal items, nonverbal evaluative cues provide useful information to the respondent in choosing the "right" answer. But for behavioural items, it is nonverbal attentiveness cues which appear to provide the more important influence on a respondent's answer. They motivate the respondent to answer more carefully and thoroughly when he or she is asked to recall information, even if the information is detrimental to the respondent's self presentation.

These same motivating attentiveness cues are responsible for mode differences in behaviour/recall items devoid of social desirability, and they also appear to cause mode differences in nonresponse patterns. Thus, the lack of any mode difference in nonresponse for attitudes argues for the irrelevance of nonverbal cues for decisions about whether to answer,

unless the subject of the question is a behaviour.

It may be that the motivating attentiveness cues could also cause mode differences in behaviour/recall items at the very socially desirable and very socially undesirable levels if the respondent is asked to provide a great deal of detail. The variables here do not test this possibility, but it would be necessary to do so to understand whether there can be mode differences on behavioural questions at the extreme ends of the desirability continuum. Such a test would also clarify whether it is attentiveness cues which always cause mode differences on behavioural questions or whether social norms (item characteristic desirability) really are more important at these extremes for both attitude and behaviour questions, as appears to be the case in these data.

Where an attitudinal item is controversial, that is, where giving a particular response could be desirable to one subgroup of respondents but undesirable to another group, answers given in personal interviews will tend to reflect the influence of interviewer characteristics. Thus, it would be difficult to determine whether some nonverbal cues take precedence over others in guiding responses. Evaluative cues linked to item characteristic desirability may not be so readily conveyed by the interviewer on these topics (due to his or her own indecision) or they may not have the expected influence because of conflicts with the interviewee's opinion. The respondent, thus, may rely more on attentiveness cues indicating affiliation in these situations and frame responses to encourage greater production of these. Or, perhaps, they simply rely more on stereotypes related to what they conceive a desirable answer would be to an interviewer having a particular type of appearance. In either case, the result would be the absence of a visible mode effect.

One further necessary modification to the theoretical ideas presented in Chapter 2 is to place more importance on nonverbal comprehension cues. These will result in mode differences on complex questions despite the degree of social desirability involved in giving a particular answer. Thus, where mode differences have appeared and it seems that the telephone respondents are giving either more or less socially desirable answers than face-to-face respondents, it is imperative to rule out the possibility that complexity of

question format is the reason for the mode difference.

There is a danger in generalizing to such a great extent from so few examples. These theoretical modifications should obviously be tested more broadly in the interest of pinpointing how mode differences originate. Doing so would require a critical reappraisal of the validity of ratings made of the desirability or undesirability of answering in a particular manner. Thus, in this study, it was suggested that the social desirability of a response for some items was not necessarily the same for all sub-groups of respondents. Furthermore, the very meaning of an item and, thus, its desirability could differ for various respondent subgroups. The possibly confounding role that grouping of several items together in one section of the questionnaire could play on the respondent's concern (or lack of it) for self-presentation in his/her manner of responding was also mentioned. All these factors would have to be taken into account in choosing the best items for testing these modified hypotheses in future research.

B. Place in the Literature

This study represents an attempt to introduce some order into the morass of contradictory findings on mode effects and "sensitive questions" by insisting that we have to think about why there should or should not be mode differences. Furthermore, any rationale for such mode differences must consider the nature of nonverbal communication cues which are present in face-to-face interaction. Explanations must also consider how such cues might be expected to relate to the content of the questionnaire items themselves in terms of their social desirability, because such desirability is usually at the basis of an item's "sensitivity". From these two underlying ideas, some hypotheses emerged which were tested on data from the 1988 Edmonton Area Survey. Results indicate some support for the ideas presented, but there were enough contrary findings to point towards the existence of a greater complexity than expected.

Most studies on mode differences which do reflect on the role that nonverbal communication must be playing simply state that it is known to perform various

communicative functions. But they do not go further to elaborate on what sorts of mode differences should, then, be expected because of these functions (e.g., Sykes and Collins, 1988). Or, they may describe some differences which are likely based on a partial analysis of nonverbal feedback functions, as when it is thought that more difficult cognitive tasks, such as recalling, will be vulnerable to media effects. Problems in this case are seen to arise because of the faster pace of telephone interviews which, in turn, is due to a need for verbal flow to compensate for the lack of other feedback cues (Sykes and Collins, 1988). In other words, more talking is necessary without nonverbal cues and, therefore, the respondent cannot think as well. The specifics of how nonverbal cue functions should create mode differences are not considered. This is probably because the cues are not looked at in depth, in relation to how they could affect communication as it would interrelate with particular items due to their content.

This study has tried to argue a reasonable case for why specific nonverbal cue functions should make a difference on particular types of items, so that we should be able to predict where certain differences in response patterns will occur. The study has been only partially successful in this task, but there were enough unanticipated problems with many of the variables chosen that a fair test of the hypotheses may not have been possible. To mention just a few of the complicating factors, some items had desirable responses for some people but not others, at least one item may have been difficult to understand, and there was some question about whether some items were truly neutral, in terms of social desirability.

Another innovative feature of the study's design was that it attempted to demonstrate that the degree of desirability of an item's response was a necessary ingredient in accounting for mode differences. A somewhat similar idea was tested in the previously mentioned study cited in Bradburn and Sudman (1979) which looked at the way degree of threat interacted with the method of questionnaire administration in producing distortion. That study, though, was limited to examining behavioural items, whereas the present study included attitudinal questions. Furthermore, the Bradburn and Sudman study did not attempt to establish the ordering of threat to the respondents in the items, whereas this study did attempt an

assessment of degree of desirability. The Bradburn and Sudman study used different populations for each question, while the same population was used throughout here. Finally, the Bradburn and Sudman work was measuring each method's deviation from a true, known value for each topic, while this study looked at relative deviations of values of each method from the other on items with unknown "true" values.

There were some areas where the present study replicated previous findings about mode differences, though in most cases, these findings had not necessarily been tied to either social desirability or "sensitivity". Thus, the findings of a basic dichotomy between behavioural and attitudinal items in terms of the effect of mode (Sudman and Bradburn, 1974) was replicated here. The earlier study stated that the response effect of telephone administration was greater for attitudinal items, whereas looking at results from a social desirability viewpoint, I would phrase the finding as a greater effect of face-to-face administration on social desirability of response for attitudinal items (when particular responses are considered somewhat socially desirable or undesirable). Nonetheless, the basic finding of a greater effect of mode on attitudinal items was reproduced here.

What is interesting, though, is that if one looks at a different measure of mode effect in the present study, nonresponse, mode seems to have more of an effect on behavioural items. In other words, there is significantly less nonresponse face-to-face for behavioural but not for attitudinal items. That nonresponse may differ by mode according to whether behavioural or attitudinal questions are examined is a finding not previously reported. This result opens up speculation that different nonverbal cues might be more important for each type of item (e.g., evaluative cues for attitudes and attention-showing motivational cues for behaviours). Such an idea would have to be tested experimentally, however.

Another replication involved Sudman and Bradburn's (1974) finding that response effects decrease as an item's saliency increases. If the fact that items have responses which are either very socially desirable or undesirable makes such items salient, then this study offers evidence that the effect of mode is not important on these "salient" questions, that is, if mean or proportion differences are the measures of mode effects. Saliency (in this sense) of

the item is not important, though, if nonresponse or variability are used as the measures of mode effects. Thus, this study may provide evidence that saliency's relationship to mode effects must be specified by the type of measurement of mode effects.

Although it was not a specific part of the hypotheses, there was evidence for replication of previous findings of greater difficulties in communicating complex questions over the telephone. This difficulty was shown by the apparently more undesirable responses of telephone respondents on the impaired driver punishments question.

It was stated at the beginning that one of the greatest concerns behind the whole idea of carrying out mode comparison studies is to find out whether one mode of questionnaire admininstration produces more error than another. That problem has not really been addressed in this thesis because a) attitudes (which do not have a right or wrong answer) were examined, and b) the answers to behavioural questions were not verified. If face-to-face responses seemed more socially desirable when attitudes were the subject, there is no way of knowing whether these or the less socially desirable telephone answers were truer to the respondents' "real" attitudes. Despite the fact that error attributable to mode could not be measured here, it is still a useful exercise to attempt to understand why and where divergences between modes will appear.

The fact that there appeared to be so much better recall on face-to-face responses, and especially when this recall factor was so strong that it seemed to override the social desirability factor, might be a more worrisome concern to survey researchers who want to use telephone administration methods. Mode does seem to introduce a kind of error, of omission, on these items and it may be worthwhile to concentrate efforts on ways to increase the level of motivation for telephone respondents so that their recall is improved.

Apparently there are ways to help eliminate differences in recall between face-to-face and telephone samples. For example, Colombotos (1969) obtained much more thorough responses over the phone when he asked physicians to list journals they read instead of merely asking for the number of publications read. Another study on memory of medical events obtained similar responses face-to-face and by telephone (Sudman *et al.*, 1984). It is possible

that the use of bounded recall procedures here produced an improvement in telephone respondents' memories. Further investigations should be made into whether such procedures are applicable to any sort of subject matter, as well as how to adapt them to particular question formats.

This study has provided some suggestions that the kinds of mode differences discussed here are consistent with those one would expect if nonverbal communication cues were responsible. But there is no way to prove with these data that such cues are actually causing these differences without carrying out controlled experiments. If the same questionnaire item is administered with and without a particular nonverbal cue by the same person in the same setting, then we can be more certain that the cue itself is responsible for differences obtained. In the present case, we cannot even be entirely certain that these cues actually occurred. Even assuming that they did occur in the face-to-face interviews, not enough is known yet about the relative strength of each type of cue in comparison to the others, or about the range of applicability of each type of cue, to be able to feel confident in making predictions about their effects. This study, then, can only be seen as an exploratory first step which has uncovered some of the potential problem areas.

C. Directions for Further Research

It is obvious that nonverbal communication cues have to be at the basis of most differences between telephone and face-to-face survey responses. Therefore, as just stated above, there needs to be some experimental work done on such communication to clarify its role in influencing how people answer questions. What precisely is communicated by a gesture? Are there factors that make its impact greater? Which cues take precedence in which situations? Can these cues be eliminated? These are all issues which would be of importance in understanding why a different answer might be given in personal and telephone interviews.

But there is also some evidence that what happens when communication is mediated by a telephone is not simply communication minus a nonverbal component but, possibly, a very different kind of communication itself. Concentration on what one hears alone seems to

create its own set of differences, such as greater effects due to quality of voice, the strength of an argument assuming more importance, and so on. Thus, improved understanding of these "audio" effects is also likely to be necessary for real comprehension of bases for mode differences. There is an obvious need for further experimental research in this area as well as in the nonverbal area. Mode differences must be seen as resulting from a combination of nonverbal communication effects and audio-focused communication effects.

Probably the most easily accomplished goals in furthering research on social desirability-related mode differences lie in the direction of improving the ability of variables to measure degree of social desirability. Obviously, the social desirability ratings of particular responses to questionnaire items should first be tested on a larger and more varied sample (from the same population on which the survey is carried out) in order to eliminate some of the biases which were apparent here. The ideal would be to carry out the social desirability assessment using the same individuals who were the respondents in the actual survey, so that we could say with more authority that a person's rating of the social desirability of a response is or is not related to the actual response he or she makes.

Second, greater care must be taken in this desirability assessment to ensure that the respondents truly have the same understanding of their task and also that items are more carefully placed in the rating questionnaire to eliminate any possibility of grouping effects on responses. This desirability questionnaire might profit from offering a greater range of response alternatives (e.g., a 1 to 7 or 9 scale on desirability rather than the five choices provided here). This would allow for the existence of possible subtler shadings in desirability ratings which may be relevant to the operation of nonverbal cues.

Third, some of the testing problems encountered here might be dispensed with by devoting more thought to the choice of fewer potentially problematic variables, possibly through more prior research into likely trouble areas such as gender-based differences in feelings about unemployment. Certainly, adding more variables to test might have compensated for some of the problems with those chosen, so a wider range of topics should be examined. As well, there must be more attention paid to ensuring that the use of variables

deriving from the same section of the questionnaire is not bringing its own set of effects. There should also be equal numbers of attitude and behaviour items within each category of social desirability.

Covering a wider range and a larger number of variables would allow the researcher to be more certain that there really was a difference in the way mode affects responses to behaviour as opposed to attitude questions and that such effects really do have some relationship to social desirability. As the results now stand, it appears that mode and the item characteristic type of social desirability interact in a fairly limited way. In other words, attitudinal items having somewhat socially desirable or undesirable responses are answered in a more socially desirable direction face-to-face. There is no difference by mode on either attitudinal or behavioural items which have very socially desirable or undesirable answers. It would seem, then, that the effect of nonverbal evaluative feedback cues on the part of the interviewer is not as strong as hypothesized, since it is limited here to attitudes, but it does seem to operate on those attitudinal items having responses which are not strongly desirable or undesirable, as hypothesized.

On the other hand, it appears that nonverbal cues which convey attention on the interviewer's part and arouse motivation to answer may play a more important role than had been hypothesized. Thus, respondents did not seem to be trying to present a better image of themselves face-to-face on behavioural items having somewhat socially undesirable responses. Instead, they seemed to be making a greater effort face-to-face to recall and answer thoroughly on undesirable topics. So, motivating effects of nonverbal cues appear to play a stronger role in face-to-face situations on behavioural items than do evaluative cues.

Such conclusions obviously require further testing since they have been based on the analysis of very few variables which were also subject to the probable effects of the confounding influences detailed above. They do, however, point to some intriguing areas for further investigation, such as the possibility that certain kinds of nonverbal cues are more important than others in determining responses on particular types of items. The only way to

be certain, though, about any conclusions concerning the operation of these cues is to conduct controlled experiments which exhibit all the rigour of design recommended by Webb and Campbell (1973).

References

Aneshensel, C. S.; Frerichs, R. R.; Clark, V. A.; and Yokopenic, P. A.
1982 Telephone Versus In-Person Surveys and Community Health Status. *American Journal of Public Health*, 72:1071-1102.

Argyle, M.
1969 *Social Interaction*. London, Methuen.

Argyle, M. and Dean, J.
1965 Eye-Contact, Distance and Affiliation. *Sociometry*, 28:289-304.

Aronson, S.
1971 The Sociology of the Telephone. *International Journal of Comparative Sociology*, XII:154-158.

Aronson, S. H.
1977 Bell's Electrical Toy: What's the Use? The Sociology of Early Telephone Usage. In I. de Sola Pool (ed.), *The Social Impact of the Telephone*. Cambridge, the MIT Press, pp. 15-39.

Ball, D. W.
1968 Toward a Sociology of Telephones and Telephoners. In M. Truzzi (ed.), *Sociology and Everyday Life*. Englewood Cliffs, N. J., pp. 59-74.

Ball, P.; Giles, H.; and Hewstone, M.
1985 Interpersonal Accommodation and Situational Construals: an Integrative Formalisation. In H. Giles and R. N. St.Clair (eds.), *Recent Advances in Language, Communication, and Social Psychology*. London, Lawrence Erlbaum Associates Ltd., pp. 263-286.

Biemer, P. P.
1988 Measuring Data Quality. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls II, J. Waksberg (eds.), *Telephone Survey Methodology*. New York, John Wiley and Sons, pp. 273-282.

Birdwhistell, R. L.
1970 *Kinesics and Context*. Philadelphia, University of Pennsylvania Press.

Bishop, G. F.; Tuchfarber, A. J.; and Oldendick, W.
1986 Opinions on Fictitious Issues: the Pressure to Answer Survey Questions. *Public Opinion Quarterly*, 50:240-250.

Boettinger, H. M.
1976 *The Telephone Book*. Croton-on-Hudson, Riverwood Press.

Bradburn, N. M.
1983 Response Effects. In P. H. Rossi, J. D. Wright, and A. Anderson (eds.), *Handbook of Survey Research*. Orlando, Academic Press, Inc., pp.289-328.

Bradburn, N. M. and Sudman, S.
1979 *Improving Interview Method and Questionnaire Design. Response Effects to Threatening Questions in Survey Research*. San Francisco, Jossey-Bass Publishers.

Brown, B. L. and Bradshaw, J. M.
1985 Towards a Social Psychology of Voice Variations. In H. Giles and R. N. St.Clair (eds.), *Recent Advances in Language, Communication, and Social Psychology*. London, Lawrence Erlbaum Associates Ltd., pp. 144-181.

Cannell, C. F.; Oksenberg, L.; and Converse, J. M. (eds.)
1979 *Experiments in Interviewing Techniques. Field Experiments in Health Reporting, 1971-1977*. Research Report Series, Institute for Social Research, Survey Research Center, University of Michigan.

Colombotos, J.
1969 Personal Versus Telephone Interviews: Effects on Responses. *Public Health Reports*, 84:773-782.

Coombs, L. and Freedman, R.
1964 Use of Telephone Interviews in a Longitudinal Fertility Study. *Public Opinion Quarterly*, 28:112-117.

Crowne, D. and Marlowe, D.
1964 *The Approval Motive*. New York, John Wiley and Sons.

deLamater, J.
1982 Response Effects of Question Content. In W. Dijkstra and J. van der Zouwen (eds.), *Response Behaviour in the Survey Interview*. London, Academic Press, pp. 13-48.

deLeeuw, E. D. and van der Zouwen, J.
1988 Data Quality in Telephone and Face-to-Face Surveys: a Comparative Meta-analysis. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls II, J. Waksberg (eds.), *Telephone Survey Methodology*. New York, John Wiley and Sons, pp. 283-299.

De Maio, T.
1984 Social Desirability and Survey Measurement: a Review. In C. F. Turner and E. Martin (eds.), *Surveying Subjective Phenomena*, Vol 2, New York. Russell Sage Foundation, pp. 257-281.

Dijkstra, W.
1987 Interviewing Style and Respondent Behavior. An Experimental Study of the Survey-Interview. *Sociological Methods and Research*, 16:309-334.

Edwards, A. L.
1957 *The Social Desirability Variable in Personality Assessment and Research*. New York, Dryden Press.

Fowler, F. J.
1984 *Survey Research Methodology*. Beverly Hills, Sage Publications.

Frey, J. H.
1983 *Survey Research by Telephone*. Beverly Hills, Sage Publications.

Giles, H.
1980 Accommodation Theory: Some New Directions. In M. W. S. deSilva (ed.), *Aspects of Linguistic Behaviour. Festschrift for R. B. LePage*. York Papers in Linguistics, No. 9, York, Universit of York.

Giles, H. and Smith, P.
1979 Accommodation Theory: Optimal Levels of Convergence. In H. Giles and R. N. St. Clair (eds.), *Language and Social Psychology*. Oxford, Basil Blackwell, pp. 45-65.

Goffman, E.
1981 *Forms of Talk*. Philadelphia, University of Pennsylvania Publications in Conduct and Communication.

Goodman, G. S.
1982 *Winning by Telephone. Telephone Effectiveness for Business Professionals and Consumers*. Englewood Cliffs, New Jersey, Prentice-Hall, Inc.

Gove, W. R. and Geerken, M. R.
1977 Response Bias in Surveys of Mental Health: an Empirical Investigation. *American Journal of Sociology*, 82:1289-1317.

Groves, R. M.
1987 Research on Survey Data Quality. *Public Opinion Quarterly*, 51:S156-S172.

Groves, R. M. and Kahn, R. L.
1979 *Surveys by Telephone. A National Comparison with Personal Interviews*. New York, Academic Press.

Henson, R.; Cannell, C. F.; and Roth, A.
1978 Effects of Interview Mode on Reporting of Moods, Symptoms, and Need for Social Approval. *The Journal of Social Psychology*, 105:123-129.

Henson, R.; Cannell, C. F.; and Lawson, S.
1979 An Experiment in Interviewer Style and Questionnaire Form. In C. Cannell L. Oksenberg, and J. M. Converse (eds.), *Experiments in Interviewing Tehniques. Field Experiments in Health Reporting, 1971-1977*. Research Report Series, Institute for Social Research, University of Michigan, pp. 21-44.

Herman, J. B.
1977 Mixed Mode Data Collection: Telephone and Personal Interviewing. *Journal of Applied Psychology*, 62:399-404.

Hochstim, J. R.
1967 A Critical Comparison of Three Strategies of Collecting Data From Households. *Journal of the American Statistical Association*, 62:976-989.

Janofsky, A. I.
1971 Affective Self-Disclosure in Telephone Versus Face-to-Face Interviews. *Journal of Humanistic Psychology*, 11:93-103.

Jordan, L. A.; Marcus, A. C.: and Reeder, L. G.
1980 Response Styles in Telephone and Household Interviewing: a Field Experiment. *Public Opinion Quarterly*, 44:210-222.

Keller, S.
1977 The Telephone in New (and Old) Communities. In I. de Sola Pool (ed.), *The Social Impact of the Telephone*. Cambridge, The MIT Press, pp. 281-198.

Kinzel, C. and Adebayo, A.
1988 All Alberta Study 1988, Sampling Report. Edmonton Area Series Report No. 58. Population Research Laboratory, Department of Sociology, University of Alberta.

Klecka, W. R. and Tuchfarber, A. J.
1978 Random Digit Dialing: a Comparison to Personal Surveys. *Public Opinion Quarterly*, 42:105-114.

Körmendi, E.
1988 The Quality of Income Information in Telephone and Face-to-Face Surveys. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls II, J. Waksberg (eds.), *Telephone Survey Methodology*. New York, John Wiley and Sons, pp. 341-356.

Lalu, N. M.
1988 Telephone and Face-to-Face Interviewing: a Comparison. Edmonton Area Series Report No. 64. Population Research Laboratory, Department of Sociology, University of Alberta.

Larsen, O. N.
1952 The Comparative Validity of Telephone and Face-to-Face Interviews in the Measurement of Message Diffusion From Leaflets. *American Sociological Review*, 17:471-476.

Mangione, T. W.; Hingson, R.; and Barrett, J.
1982 Collecting Sensitive Data. *Sociological Methods and Research*, 10:337-346.

Matarazzo, J. D.; Wiens, A. N.; and Saslow, G.
1965 Studies in Interview Speech Behavior. In L. Krasner and L. Ullman (eds.), *Research in Behavior Modification. New Developments and Implications*. New York, Holt, Rinehart, and Winston, Inc., pp. 179-210.

Matarazzo, J. D. and Wiens, A. N.
1972 *The Interview. Research on Its Anatomy and Structure*. Chicago, Aldine Atherton.

Mayer, M.
1977 The Telephone and the Uses of Time. In I. de Sola Pool (ed.), *The Social Impact of the Telephone*. Cambridge, The MIT Press, pp. 225-245.

McGuire, W. J.
1985 Attitudes and Attitude Change. In G. Lindzey and E. Aronson (eds.), *The Handbook of Social Psychology*, Third edition, Vol. II, Chapter 19, pp. 233-346.

McLuhan, M.
1964 *Understanding Media. The Extensions of Man*. New York, McGraw-Hill Book Company.

Oksenberg, L.; Coleman, L.; and Cannell, C. F.
1986 Interviewers' Voices and Refusal Rates in Telephone Surveys. *Public Opinion Quarterly*, 50:97-111.

Oksenberg, L. and Cannell, C.
1988 Effects of Interviewer Vocal Characteristics on Nonresponse. In R. M. Groves, P. P. Biemer, L. Lyberg, J. T. Massey, W. L. Nicholls II, and J. Waksberg (eds.), *Telephone Survey Methodology*. New York, John Wiley and Sons, pp. 257-269.

Orne, M. T.
1962 On the Social Psychology of the Psychological Experiment: with Particular Reference to Demand Characteristics and Their Implications. *American Psychologist*, 17:776-783.

Phillips, D. L. and Clancy, K. J.
1970 Response Biases in Field Studies of Mental Illness. *American Sociological Review*, 35:503-515.

Phillips, D. L., and Clancy, K. J.
1972 Some Effects of Social Desirability in Survey Studies. *American Journal of Sociology*, 77:921-940.

Reid, A. A. L.
1977 Comparing Telephone with Face-to-Face Contact. In I. de Sola Pool (ed.), *The Social Impact of the Telephone*. Cambridge, The MIT Press, pp. 386-414.

Rogers, T. F.
1976 Interviews by Telephone and In Person: Quality of Responses and Field Performance. *Public Opinion Quarterly*, 40:51-65.

Scherer, K. R.
1980 The Functions of Nonverbal Signs in Conversation. In R. N. St.Clair and H. Giles (eds.), *The Social and Psychological Contexts of Language*. Hillsdale, New Jersey, Lawrence Erlbaum Associates, pp. 225-244.

Schuessler, K.; Hittle, D.: and Cardascia, J.
1978 Measuring Responding Desirably with Attitude-Opinion Items. *Social Psychology*, 41:224-235.

Sebastian, R. J. and Ryan, E. B.
1985 Speech Cues and Social Evaluation: Markers of Ethnicity, Social Class, and Age. In H. Giles and R. N. St. Clair (eds.), *Recent Advances in Language, Communication, and Social Psychology*. London, Lawrence Erlbaum Associates, Ltd., pp. 112-143.

Selltiz, C.; Wrightsman, L. S.; and Cook, S. W.
1976 *Research Methods in Social Relations*, 3rd edition. New York, Holt, Rinehart and Winston.

Short, J.; Williams, E.; and Christie, B.
1976 *The Social Psychology of Telecommunications*. London, John Wiley and Sons.

Siegman, A. W.
1980 Interpersonal Attraction and Verbal Behavior in the Initial Interview. In R. N. St.Clair and H. Giles (eds.), *The Social and Psychological Contexts of Language*. Hillsdale, New Jersey, Lawrence Erlbaum Associates, pp. 73-99.

Sillars, A. L.
1982 Attribution and Communication. Are People 'Naive Scientists' or Just Naive? In M. E. Roloff and Charles R. Berger (eds.), *Social Cognition and Communication*. Beverly Hills, Sage Publications, pp. 73-106.

Smith, M. D.
1985 *Woman Abuse: the Case for Surveys by Telephone*. Report No. 12, Institute for Social Research, York University.

Smith, M. D.
1986 On the Reporting of Woman Abuse: a Telephone Survey Experiment. Toronto, Institute for Social Research, York University.

Smith, P. M.; Giles, H.; and Hewstone, M.
1980 Sociolinguistics: a Social Psychological Perspective. In R. N. St.Clair and H. Giles (eds.), *The Social and Psychological Contexts of Language*. Hillsdale, New Jersey, Lawrence Erlbaum Associates, pp. 283-298.

Smith, T. W.
1984 A Comparison of Telephone and Personal Interviewing. Chicago, National Opinion Research Center.

Street, R. L., jr. and Giles, H.
1982 Speech Accommodation Theory. A Social Cognitive Approach to Language and Speech Behavior. In M. E. Roloff and C. R. Berger (eds.). *Social Cognition and Communication*. Beverly Hills, Sage Publications, pp. 193-226.

Stricker, L. J.
1963 Acquiescence and Social Desirability Response Styles, Item Characteristics, and Conformity. *Psychological Reports*, 12:319-341.

Sudman, S.; Greeley, A.; and Pinto, L.
1965 The Effectiveness of Self-Administered Questionnaires. *Journal of Marketing Research*, 2:293-297.

Sudman, S. and Bradburn, N. M.
1974 *Response Effects in Surveys. A Review and Synthesis*. Chicago, Aldine Publishing Company.

Sudman, S.; Bradburn, N.; Blair, E.; and Stocking, C.
1977 Modest Expectations: the Effects of Interviewers' Prior Expectations on Response. *Sociological Methods and Research*. 6:177-182.

Sudman, S.; Finn, A.; and Lannon, L.
1984 The Use of Bounded Recall Procedures in Single Interviews. *Public Opinion Quarterly*, 48:520-524.

Sykes, W. M. and Collins, M.
1987 Telephone and Face-to-Face Interviewing in the United Kingdom. *Survey Methodology*, 13:15-28.

Sykes, W. and Collins, M.
1988 Effects of Mode of Interview: Experiments in the UK. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls III, and J. Waksberg (eds.), *Telephone Survey Methodology*. New York, John Wiley and Sons, pp. 301-320.

Walsh, A.
1987 Teaching Understanding of Logit Regression. *Teaching Sociology*, 15:178-183.

Webb, E. J. and Campbell, D. T.
1973 Experiments on Communication Effects. In I. de Sola Pool and F. W. Frey (eds.), *Handbook of Communication*. Chicago, Rand McNally College Publishing Company, pp. 938-952.

Weiss, C. H.
1968 Validity of Welfare Mothers' Interview Responses. *Public Opinion Quarterly*, 32:622-633.

Wildman, R. C.
1977 Effect of Anonymity and Social Setting on Survey Responses. *Public Opinion Quarterly*, 41:74-79.

Williams, E.
1975 Medium or Message: Communications Medium as a Determinant of Interpersonal Evaluation. *Sociometry*, 38:119-130.

Williams, E.
1977 Experimental Comparisons of Face-to-Face and Mediated Communication. *Psychological Bulletin*, 84:963-976.

Appendix One

An Assessment of Social Desirability

The following questions ask people about behaviours and attitudes that might be considered good or bad by most people in our society. Please rate the response provided for each question on the following scales of social desirability. In other words, how socially desirable do you feel it would be to give that response to that question?

(1) Very socially desirable

(2) Fairly socially desirable

(3) Social desirability would not be a factor in how a person answers this question

(4) Somewhat socially undesirable

(5) Very socially undesirable

(9) A question might be classified as "9" because some respondents would feel the socially desirable response lies in one direction while many respondents feel it lies in the opposite direction. [The (9) rating was only to be provided on the experts' questionnaire]

Rating	Question
	There should be stricter penalties for parents who do not properly secure their children in car safety seats.
( )	Response = Agree
	How often do you get together with your friends?
( )	Response = Often
	Punishments for impaired drivers are becoming too severe.
( )	Response = Agree
	All in all how satisfied with life are you these days?
( )	Response = Satisfied
	To get ahead in life you really need a university education.
( )	Response = Agree
	In the past three years (or as long as you have lived in Alberta, if less than three years), have you or any other members of your household received Welfare assistance of any kind?
( )	Response = Yes
	It would be better for a pre-school child's emotional development if the mother didn't work outside the home.
( )	Response = Agree.
	[ACCESS Network] Are you aware of the television broadcast service?
( )	Response = Yes

I don't think the Federal government cares much about what people like me think.

( ) Response = Agree

How many times did you attend church in the last four weeks?

( ) Response = Often

In the past three years (or as long as you have lived in Alberta, if less than three years), have you or any other members of your household made any contributions to a Food Bank?

( ) Response = Yes

How often do you get together with relatives?

( ) Response = Often

About how many hours in the past week did you watch ACCESS Television on cable channel 9?

( ) Response = Many Hours

In the past three years (or as long as you have lived in Alberta, if less than three years), have you or any other members of your household drawn any unemployment insurance (UIC)?

( ) Response = Yes

The research done in universities benefits society.

( ) Response = Agree

It is important to allow women to become ordained ministers/priests.

( ) Response = Agree

About how many hours in the past week have you personally watched the American Broadcasting Service (PBS) KSPS from Spokane on cable channel 13?

( ) Response = Many Hours

How many groups and organizations do you belong to, including professional, union, recreational and church groups, etc.?

( ) Response = Many Groups

There is a good chance that someone in my household will be unemployed in the next year.

( ) Response = Agree

In the next two years do you or anyone else in your household plan to take a university-credit course at an Alberta college?

( ) Response = Yes

People with high incomes should pay a greater share of the taxes than they do now.

( ) Response = Agree

Now looking ahead - do you think that a year from now you (and your family), will be better off financially, or worse off, or just about the same as now?

( ) Response = Better Off

About how many hours in the past week have you personally watched television?

( ) Response = Many Hours

Federal and Provincial Rights Codes make it illegal to discriminate in employment, housing and services because of race, color, religion, age, sex and so on. Human Rights Commissions have recommended that legislation be passed making it illegal to discriminate on the basis of sexual orientation; that is, heterosexuality, homosexuality or bisexuality. How strongly do you agree or disagree with this recommendation?

( ) Response = Agree

In the past three years (or as long as you have lived in Alberta, if less than three years), have you personally been unemployed for up to a month?

( ) Response = Yes

It's okay to drive soon after having two or three alcoholic drinks (within 1-2 hours).

( ) Response = Agree

In total, how many years of schooling do you have? This includes the total of grade school, high school, vocational, technical, and university?

( ) Response = Many Years

In the past three years (or as long as you have lived in Alberta if less than three years), have you or any other members of your household received financial help from relatives?

( ) Response = Yes

Appendix Two

Results of Social Desirability Assessment

Variable		Rating* and N of Each					
		1	2	3	4	5	9
1. There should be stricter penalties for parents who do not properly secure their children in car safety seats							
Agree	Experts	2	5				1
	Interviewers	5	1		2		
	Friends	10	5				1**
	Totals	17	11		2		2
2. How often do you get together with friends?							
Often	Experts	1	3	4			
	Interviewers	3	2	5			
	Friends	6	5	4			1
	Totals	10	10	13			1
3. Punishments for impaired drivers are becoming too severe.							
Agree	Experts				3	5	
	Interviewers	3	1	1	3	2	
	Friends	2	2	2	5	5	
	Totals	5	3	3	11	12	
4. All in all how satsified with life are you these days?							
Satisfied	Experts		6	2			
	Interviewers	4	3	3			
	Friends	2	7	5	2		
	Totals	6	16	10	2		
5. To get ahead in life you really need a university education.							
Agree	Experts	1	1	4			2
	Interviewers	2	3	3			2
	Friends	1	7	5	2		1
	Totals	4	11	12	2		5

* It will be recalled that a rating of 1 meant very socially desirable, 2, somewhat socially desirable, 3, neutral in social desirability, 4, somewhat socially undesirable, 5, very socially undesirable, and 9 meant socially desirable for some, but not for others.

** A few friends and interviewers were mistakenly given questionnaires with a (9) rating option. Only experts were to have been provided with this choice.

		Rating and N of Each					
		1	2	3	4	5	9
6. In the past three years (or as long as you have lived in Alberta, if less than three years), have you or any other members of your household received Welfare assistance of any kind? Yes	Experts				2	6	
	Interviewers	2	1	2	4	1	
	Friends		2	2	5	7	
	Totals	2	3	4	11	14	
7. It would be better for a pre-school child's emotional development if the mother didn't work outside the home. Agree	Experts		3	1			4
	Interviewers	1	6	1			2
	Friends	3	6	2	2	1	2
	Totals	4	15	4	2	1	8
8. [ACCESS Network] Are you aware of the television broadcast service? Yes	Experts		5	3			
	Interviewers	2		8			
	Friends	6	5	5			
	Totals	8	10	16			
9. I don't think the Federal government cares much about what people like me think. Agree	Experts		2	5	1		
	Interviewers	1	3	4	2		
	Friends	2	3	7	2	2	
	Totals	3	8	16	5	2	
10. How many times did you attend church in the last four weeks? Often	Experts		4	3			1
	Interviewers	2	2	6			
	Friends	2	5	6	1	1	1
	Totals	4	11	15	1	1	2
11. In the past three years (or as long as you have lived in Alberta, if less than three years), have you or any other members of your household made any contributions to a Food Bank? Yes	Experts	3	5				
	Interviewers	4	3	1	2		
	Friends	8	2	4	1	1	
	Totals	15	10	5	3	1	

	Rating and N of Each					
	1	2	3	4	5	9
12. How often do you get together with relatives? Often						
Experts		2	6			
Interviewers	2	4	4			
Friends	3	11	2			
Totals	5	17	12			
13. About how many hours in the past week have you personally watched ACCESS Network Television on cable channel 9? Many Hours						
Experts		2	6			
Interviewers	1	1	8			
Friends	2	5	5	2	1	1
Totals	3	8	19	2	1	1
14. The research done in universities benefits society. Agree						
Experts	1	3	4			
Interviewers	1	5	3			1
Friends	5	5	4			2
Totals	7	13	11			3
15. It is important to allow women to become ordained ministers/priests. Agree						
Experts		1	4			3
Interviewers	2	3	4			1
Friends	5	5	3	1		2
Totals	7	9	11	1		6
16. About how many hours in the past week have you personally watched the American Broadcasting Service (PBS) from Spokane on cable channel 13? Many Hours						
Experts		4	4			
Interviewers	1	1	8			
Friends	2	7	4	1	1	1
Totals	3	12	16	1	1	1
17. How many groups and organizations do you belong to, including professional, union, recreational and church groups, etc.? Many Groups						
Experts		6	2			
Interviewers	1	5	4			
Friends	3	9	3	1		
Totals	4	20	9	1		

	Rating and N of Each					
	1	2	3	4	5	9
18. There is a good chance that someone in my household will be unemployed next year. Agree						
Experts			2	6		
Interviewers	1	1	5	2	1	
Friends	1		5	7	3	
Totals	2	1	12	15	4	
19. In the next two years do you or anyone else in your household plan to take a university-credit course at an Alberta college? Yes						
Experts		4	4			
Interviewers	2	3	5			
Friends	3	11	2			
Total	5	18	11			
20. People with high incomes should pay a greater share of the taxes than they do now. Agree						
Experts	1	2	4			
Interviewers	2	4	3	1		
Friends	2	6	3	2	2	1
Totals	5	12	10	3	2	1
21. Now looking ahead - do you think that a year from now you (and your family) will be better off financially, or worse off, or just about the same as now? Better Off						
Experts		4	4			
Interviewers	3	5	2			
Friends	4	7	5			
Total	7	16	11			
22. About how many hours in the <u>past week</u> have <u>you</u> personally watched television? Many Hours						
Experts	1		3	3		1
Interviewers	1	2	3	4		
Friends	1	3	1	9	2	
Totals	3	5	7	16	2	1

		Rating and N of Each					
		1	2	3	4	5	9
23. Federal and Provincial Human Rights Codes make it illegal to discriminate in employment, housing and services because of race, color, religion, age, sex and so on. Human Rights Commissions have recommended that legislation be passed making it illegal to discriminate on the basis of sexual orientation; that is, heterosexuality, homosexuality or bisexuality. How strongly do you agree or disagree with this recommendation? Agree	Experts	1	3	2			2
	Interviewers	5	4		1		
	Friends	6	6	2	2		
	Totals	12	13	4	3		2
24. In the <u>past three years</u> (or as long as you have lived in Alberta, if less than three years), have you <u>personally</u> been unemployed for up to a month? Yes	Experts		1	2	5		
	Interviewers	1	2	3	4		
	Friends	2		3	10	1	
	Totals	3	3	8	19	1	
25. It's okay to drive soon after having two or three alcoholic drinks (within 1-2 hours). Agree	Experts				3	4	1
	Interviewers	2	2		1	4	1
	Friends		1	2	9	5	
	Totals	2	2	2	13	13	2
26. In total, how many years of schooling do you have? This includes total of grade school, high school, vocational, technical, and university? Many Years	Experts	1	5	1		1	
	Interviewers	3	4	2	1		
	Friends	4	9	2			
	Totals	8	18	5	1	1	
27. In the <u>past three years</u> (or as long as you have lived in Alberta, if less than three years), have <u>you or any other members of your household</u> received financial help from relatives? Yes	Experts		1	3	3		1
	Interviewers	1	2	2	5		
	Friends		1	2	11	2	
	Totals	1	4	7	19	2	1